mpr

A Mathematica Company

AN ASSESSMENT OF ALTERNATIVE COMPARISON GROUP METHODOLOGIES FOR EVALUATING EMPLOYMENT AND TRAINING PROGRAMS

FINAL REPORT

JUNE 23, 1984

HD
5715.5
.F73
A87
1984

). Box 2393
nceton, New Jersey 08540
)9) 799-2600

600 Maryland Avenue S.W.
Suite 550
Washington, D.C. 20024
(202) 484-9220

905 University Avenue
Room 203
Madison Wisconsin 53715
(608) 255-1900

HD
5715.5
.F73
A87
1984

Authors: Thomas Fraker
Rebecca Maynard
with
Lyle Nelson

Programmer: Douglas Birdzell

AN ASSESSMENT OF ALTERNATIVE COMPARISON GROUP METHODOLOGIES FOR EVALUATING EMPLOYMENT AND TRAINING PROGRAMS

FINAL REPORT

JUNE 23, 1984

Prepared by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, New Jersey 08540

Prepared for:

U.S. Department of Labor
Employment and Training
Administration
601 D Street, N.W.
Washington, D.C. 20213

Project Officer: Mr. Dan Ryan

This report was prepared for the Employment and Training Administration, U.S. Department of Labor, under Research and Evaluation Contract No. 20-11-82-15. Since contractors conducting research and evaluation reports under government sponsorship are encouraged to express their own judgments freely, this report does not necessarily represent the official opinion or policy of the Department of Labor. Mathematica Policy Research, Inc., is solely responsible for the contents of this report.

50272-101

REPORT DOCUMENTATION PAGE	1. REPORT NO.	2.	3. Recipient's Accession No.
4. Title and Subtitle: An Assessment of Alternative Comparison Group Methodologies for Evaluating Employment and Training Programs			5. Report Date: June 1984 6.
7. Author(s): Thomas Fraker and Rebecca Maynard with Lyle Nelson			8. Performing Organization Rept. No.: 7523-01
9. Performing Organization Name and Address: Mathematica Policy Research, Inc. P.O. Box 2393 Princeton, NJ 08540			10. Project/Task/Work Unit No. 11. Contract(C) or Grant(G) No. (C) DOL-20-11-82-15 (G)
12. Sponsoring Organization Name and Address: U. S. Department of Labor Employment and Training Administration Office of Research and Evaluation 601 D Street, N. W., Washington, D. C. 20213			13. Type of Report & Period Covered: Final Report 14.

15. Supplementary Notes

16. Abstract (Limit: 200 words)

The purpose of this study is to provide an empirical assessment of the reliability of net program-impact estimates generated using comparison groups constructed based on eligibility screens, cell matching, and statistical matching. The study is based on analyses of the national Supported Work experimental and control groups of youth and AFDC recipients, as well as on comparison groups for the Supported Work experimental samples generated from the Current Population Survey samples. The results indicate that, especially for young school drop-outs, it is not possible to generate reliable net program impact estimates using these ex-post comparison group procedures. For AFDC recipients, the qualitative results based on comparison groups tend to be correct, however, the magnitudes of the estimated impacts vary substantially. The direction and extent of bias in the net impact estimates tends not to be sensitive to either the comparison group construction procedure or to variation in the estimation model (e.g., random versus fixed-effects models).

17. Document Analysis a. Descriptors

Effectiveness, Employment, Evaluation, Government Policies, Statistical Analysis

b. Identifiers/Open-Ended Terms

Comparison Group Methodologies, Cell Matching, Statistical Matching, Youth, AFDC Recipients

c. COSATI Field/Group 5I, 14A

18. Availability Statement: Distribution is unlimited. Available from the National Technical Information Service, Springfield, VA 22151	19. Security Class (This Report): UNCLASSIFIED	21. No. of Pages
	20. Security Class (This Page): UNCLASSIFIED	22. Price

(See ANSI-Z39.18) See Instructions on Reverse OPTIONAL FORM 272 (4-77) (Formerly NTIS-35) Department of Commerce

CONTENTS

Chapter		Page
	EXECUTIVE SUMMARY	vii
	ACKNOWLEDGEMENTS	xiii
I	INTRODUCTION AND SUMMARY	1
	A. PROGRAM EVALUATION STRATEGIES	2
	1. True Control Group	2
	2. Advance Match Comparison Group	3
	3. Ex-Post Matched Comparison Group	4
	4. Internal Comparison Group	5
	B. OVERVIEW OF THIS STUDY	5
	1. The Comparison Group Construction Procedures	6
	2. The Evaluation of Alternative Comparison-Group Methods	11
	3. The Major Lessons and Conclusions From this Study	14
	C. ORGANIZATION OF THE REMAINDER OF THE REPORT	17
II	SAMPLE AND DATA	19
	A. THE SUPPORTED WORK YOUTH AND AFDC SAMPLES	19
	1. The Supported Work Demonstration and Sample Design	20
	2. The Youth Sample	22
	3. The AFDC Sample	27
	B. THE CPS SAMPLE USED TO DEVELOP COMPARISON GROUPS	30
	1. The CPS Youth Sample	33
	2. The CPS AFDC Sample	36
	C. THE COMBINED SUPPORTED WORK AND CPS DATA SETS	36
III	CONSTRUCTION OF COMPARISON GROUPS	39
	A. INTRODUCTION	39
	B. RANDOMLY SELECTED COMPARISON GROUPS	41
	1. Random Comparison Samples for Youth	42
	2. Random Comparison Samples for AFDC Recipients	49
	C. STRATIFIED MATCHING	55
	1. Stratified Matching for Youth	55
	2. Stratified Matching for AFDC Recipients	68
	D. STATISTICAL MATCHING	79
	1. Statistical Matching for Youth	80
	2. Statistical Matching for AFDC Recipients	87
IV	EVALUATION OF ALTERNATIVE COMPARISON GROUPS	95
	A. PRACTICAL CONSIDERATIONS IN ADOPTING COMPARISON GROUP METHODOLOGIES	96
	1. Comparability of Participant and Comparison Group Characteristics	96
	2. The Similarity of Behavioral Models for Participant and Comparison Group Members	98
	3. Adequacy of the CPS Sample Size	99

Contents (continued)

Chapter — Page

IV (continued)

B. THE QUALITY OF NET IMPACT ESTIMATES GENERATED USING THE COMPARISON GROUPS..... 101
1. Analytic Approach..... 102
2. Results of the Assessment of the Quality of the Net Impact Estimates..... 110

C. THE IMPLICATIONS OF THESE RESULTS FOR OTHER STUDIES THAT MUST RELY ON COMPARISON GROUPS..... 119

APPENDIX A: SUPPLEMENTARY TABLES FOR CHAPTER II..... 125

APPENDIX B: SUPPLEMENTARY TABLES FOR CHAPTER III..... 129

APPENDIX C: SUPPLEMENTARY TABLES FOR CHAPTER IV..... 139

APPENDIX D: STANDARD ERRORS OF NET IMPACT ESTIMATES..... 151

APPENDIX E: ALTERNATE STATISTICAL MATCH RESULTS..... 155

APPENDIX F: WEIGHTING PROCEDURES USED IN THE ANALYSIS..... 173

REFERENCES..... 179

TABLES

Number		Page
1	SUMMARY OF COMPARISON GROUP METHODOLOGIES AND SAMPLE SIZES	viii
2	SUMMARY OF NET IMPACT ESTIMATES USING ALTERNATIVE COMPARISON SAMPLES	xi
I.1	SUMMARY OF KEY FEATURES OF THE COMPARISON GROUP CONSTRUCTION PROCEDURES EVALUATED	8
I.2	DESCRIPTIVE SUMMARY OF EVALUATION ISSUES, ANALYTIC APPROACHES, AND HIGHLIGHTS OF RESULTS	12
II.1	ELIGIBILITY FOR ENROLLMENT IN THE SUPPORTED WORK YOUTH TARGET GROUP	23
II.2	ENROLLMENT IN THE SUPPORTED WORK YOUTH SAMPLE BY SITE AND CALENDAR QUARTER	24
II.3	CHARACTERISTICS OF THE SUPPORTED WORK YOUTH SAMPLE	25
II.4	DEMOGRAPHIC AND BACKGROUND CHARACTERISTICS FOR THE SUPPORTED WORK YOUTH, YEDPA YOUTH, AND THE U.S. POPULATION OF 16- TO 19-YEAR OLDS	26
II.5	ELIGIBILITY FOR ENROLLMENT IN THE SUPPORTED WORK AFDC TARGET GROUP	28
II.6	CHARACTERISTICS OF THE SUPPORTED WORK AFDC SAMPLE	29
II.7	PERCENT DISTRIBUTIONS OF THE AFDC AND WIN PARTICIPANT POPULATIONS AND OF THE SUPPORTED WORK AFDC SAMPLE, BY DEMOGRAPHIC CHARACTERISTICS	31
II.8	SCREENS APPLIED TO CPS FILES TO GENERATE SAMPLES ELIGIBLE FOR INCLUSION IN THE COMPARISON SAMPLES	34
II.9	CHARACTERISTICS OF THE CPS SAMPLES OF YOUTH AND AFDC RECIPIENTS	35
II.10	SUMMARY OF ANALYSIS VARIABLES	37
III.1	DERIVATION OF RANDOM COMPARISON SAMPLES FOR THE SUPPORTED WORK YOUTH SAMPLE	44
III.2	PERCENT DISTRIBUTIONS OF CASES FROM THE SUPPORTED WORK YOUTH SAMPLES AND CPS YOUTH RANDOM COMPARISON SAMPLES	46

List of Tables (continued)

Number		Page
III.3	DERIVATION OF RANDOM COMPARISON SAMPLES FOR THE SUPPORTED WORK AFDC SAMPLE	50
III.4	PERCENT DISTRIBUTION OF CASES FROM THE SUPPORTED WORK AFDC SAMPLES AND CPS AFDC RANDOM COMPARISON SAMPLES	52
III.5	ORDER IN WHICH POTENTIAL MATCH VARIABLES ENTERED STEPWISE EARNINGS REGRESSIONS FOR YOUTH	57
III.6	CATEGORIES OF VARIABLES USED IN STRATIFIED MATCHING FOR YOUTH	58
III.7	SUMMARY OF LONG LIST CELL MATCHING RESULTS FOR THE SUPPORTED WORK YOUTH EXPERIMENTAL SAMPLE	60
III.8	NUMBER OF CASES IN COLLAPSED CELLS FROM THE SUPPORTED WORK YOUTH EXPERIMENTAL SAMPLE AND THREE MATCHED CPS SAMPLES, USING THE LONG LIST OF MATCH VARIABLES	61
III.9	SUMMARY OF SHORT LIST CELL MATCHING RESULTS FOR THE SUPPORTED WORK YOUTH EXPERIMENTAL SAMPLE	63
III.10	PERCENT DISTRIBUTIONS OF CASES FROM THE SUPPORTED WORK YOUTH SAMPLES AND WEIGHTED DISTRIBUTIONS OF MATCHED CPS YOUTH CASES FROM THE SHORT LIST MATCH AND LONG LIST MATCHES A, B, AND C	64
III.11	ORDER IN WHICH POTENTIAL MATCH VARIABLES ENTERED STEPWISE EARNINGS REGRESSIONS FOR AFDC RECIPIENTS	69
III.12	CATEGORIES OF VARIABLES USED IN STRATIFIED MATCHING FOR AFDC RECIPIENTS	71
III.13	SUMMARY OF LONG LIST CELL MATCHING RESULTS FOR THE SUPPORTED WORK AFDC EXPERIMENTAL SAMPLE	72
III.14	NUMBER OF CASES IN COLLAPSED CELLS FROM THE SUPPORTED WORK AFDC SAMPLE AND A MATCHED CPS SAMPLE, USING THE LONG LISTS OF MATCH VARIABLES	73
III.15	SUMMARY OF SHORT LIST CELL MATCHING RESULTS FOR THE SUPPORTED WORK AFDC EXPERIMENTAL SAMPLE	75
III.16	PERCENT DISTRIBUTIONS OF CASES FROM THE SUPPORTED WORK AFDC SAMPLES AND WEIGHTED DISTRIBUTIONS OF MATCHED CPS AFDC CASES FROM THE SHORT LIST MATCH AND THE LONG LIST MATCH	76
III.17	DISTRIBUTIONS OF ACTUAL 1979 SSA EARNINGS AND 1979 SSA EARNINGS AS PREDICTED BY THE STATISTICAL MATCH EARNINGS EQUATION FOR THE SUPPORTED WORK YOUTH EXPERIMENTAL AND CPS STAGE 2 YOUTH SAMPLES	82

List of Tables (continued)

Number		Page
III.18	PERCENT DISTRIBUTIONS OF CASES FROM THE SUPPORTED WORK YOUTH SAMPLES AND CPS YOUTH STATISTICAL MATCH SAMPLE.....	84
III.19	DISTRIBUTIONS OF ACTUAL 1979 SSA EARNINGS AND 1979 SSA EARNINGS AS PREDICTED BY THE STATISTICAL MATCH EARNINGS EQUATION FOR THE SUPPORTED WORK AFDC EXPERIMENTAL AND CPS STAGE 2 AFDC SAMPLES.....	88
III.20	PERCENT DISTRIBUTION OF CASES FROM THE SUPPORTED WORK AFDC SAMPLES AND THE CPS AFDC STATISTICAL MATCH SAMPLE.....	90
IV.1	RESULTS OF CHOW TESTS OF THE SIMILARITY OF EARNINGS MODELS BETWEEN THE EXPERIMENTAL AND THE VARIOUS COMPARISON GROUP SAMPLES.....	100
IV.2	DEFINITIONS OF CONTROL VARIABLES USED IN THE ANALYSIS.....	106
IV.3	ESTIMATED NET IMPACTS OF SUPPORTED WORK ON THE ANNUAL EARNINGS OF YOUTH: ALTERNATIVE MODELS AND COMPARISON GROUPS.....	111
IV.4	ESTIMATED NET IMPACTS OF SUPPORTED WORK ON THE ANNUAL EARNINGS OF AFDC RECIPIENTS: ALTERNATIVE MODELS AND COMPARISON GROUPS.....	112
IV.5	NET IMPACT ESTIMATES USING GROUP VERSUS UNGROUPED DATA WITH THE BASIC EARNINGS MODEL.....	115
IV.6	NET IMPACT ESTIMATES GENERATED USING WEIGHTED VERSUS UNWEIGHTED REGRESSIONS FOR THE LONG LIST CELL MATCH COMPARISON GROUPS.....	117
IV.7	NET IMPACT ESTIMATES FOR THE SUPPORTED WORK AND CETA PARTICIPANTS BASED ON ALTERNATIVE COMPARISON GROUPS AND ESTIMATION TECHNIQUES.....	121

FIGURES

Number		Page
III.1	AVERAGE SSA EARNINGS FOR YOUTH: SUPPORTED WORK EXPERIMENTALS AND CONTROLS AND CPS RANDOM SAMPLES 1 AND 2.....	48
III.2	AVERAGE SSA EARNINGS FOR AFDC WOMEN: SUPPORTED WORK EXPERIMENTALS AND CONTROLS AND CPS RANDOM SAMPLES 1 AND 2...	54
III.3	AVERAGE SSA EARNINGS FOR YOUTH SUPPORTED WORK EXPERIMENTALS AND CONTROLS; WEIGHTED CPS MATCHES A, B, AND C, AND WEIGHTED CPS SHORT LIST MATCH.....	67
III.4	AVERAGE SSA EARNINGS FOR AFDC WOMEN: SUPPORTED WORK EXPERIMENTALS AND CONTROLS, AND WEIGHTED CPS SHORT AND LONG LIST MATCHES.....	78
III.5	AVERAGE SSA EARNINGS FOR YOUTH: SUPPORTED WORK CONTROLS AND STATISTICALLY MATCHED CASES FROM THE SUPPORTED WORK EXPERIMENTAL SAMPLE AND CPS RANDOM SAMPLE 2.....	86
III.6	AVERAGE SSA EARNINGS FOR AFDC WOMEN: SUPPORTED WORK CONTROLS AND STATISTICALLY MATCHED CASES FROM THE SUPPORTED WORK EXPERIMENTAL SAMPLE AND CPS RANDOM SAMPLE 2.....	92

EXECUTIVE SUMMARY

This report presents the findings from one of a number of studies funded under Department of Labor contracts and grants to address issues relevant to assessing the net impacts of employment and training programs using non-experimental methodologies. The purpose of this particular study is to provide an empirical assessment of the reliability of net impact estimates generated through the use of alternative comparision groups constructed by: (1) imposing simple eligibility screens; (2) cell matching; and (3) statistical matching.

Sample and Data

In conducting this study, we have used the national Supported Work demonstration as the program intervention to be assessed and have drawn our various comparison samples from the Current Population Survey samples. These choices were based on several considerations. First, the Supported Work Youth and AFDC samples are similar in important respects to CETA and JTPA youth and adult female participants, respectively. Second, there are randomly selected control groups for these Supported Work samples, which can be used to generate unbiased estimates of true program impacts. Third, the Supported Work and CPS data bases are reasonably comparable and both include measures of Social Security earnings for the period of 1972 through 1979--a period of time that is between three and five years prior to program entry and two to four years after. Finally, the Supported Work participants were enrolled in the program between April 1975 and July 1977, a period of time covered by the Continuous Longitudinal Manpower Survey (CLMS), which is the primary data base available for evaluating the CETA program and for providing necessary input into current employment policy.

The Supported Work sample consists of 1244 young school dropouts (566 experimentals and 678 controls) and 1602 female long-term recipients of AFDC (800 experimentals and 802 controls), none of whom had a child younger than six. In the two cohorts of the CPS used for constructing comparison samples (the 1976 and the 1977), there were a total of 14,084 youth, 2368 of whom had dropped out of school prior to completing high school, and 1995 AFDC recipients, 909 of whom had no young children.

Comparison Group Construction Procedures

In this study, we constructed six alternative comparison samples for each the youth and AFDC samples: two were random samples satisfying certain program eligibility criteria; two were constructed using cell matching techniques; and two were constructed using statistical matching procedures. Table 1 summarizes the basic approach and the resulting sample size for each comparison sample. In constructing all samples, only CPS

TABLE 1

SUMMARY OF COMPARISON GROUP METHODOLOGIES AND SAMPLE SIZES

	Youth		AFDC Recipients	
	Sample Construction Procedure	Number of Cases	Sample Construction Procedures	Number of Cases
Random Samples				
1. Stage 1 Screens	10 percent of cases Age 16 to 20	1424	All Female AFDC Recipients	1995
2. Stage 2 Screens	All 16 to 20 year old school dropouts	2368	All Female AFDC Recipients with no young children	909
Cell Match Samples				
3. Short List	16 to 20 year old dropouts that match Supported Work cases on: (1) Age, (2) Sex, and (3) Race	2365	Female AFDC recipients with no young children who match Supported Work cases on: (1) Education and (2) Race	909
4. Long List	(1) Sex, (2) Prior earnings, (3) Change in prior earnings, (4) Race, (4) Education, and (6) Age	871 to 1120	(1) Change in earnings, (2) Age, (3) Prior employment, (4) Prior Earnings, and (5) Race	554
Statistical Match Samples				
5. Initial Match	16 to 20 year old youth who most closely match supported work cases on predicted earnings: One-to-one match with a maximum difference of $500 in predicted earnings	558	Female AFDC recipients with no young children who most closely match Supported Work cases on predicted earnings: One-to-one match with a maximum difference of $500 in predicted earnings	582
6. Statistiscal Match Within Strata	Closest match within strata detined by sex and race. Penalties imposed for previous selection into the sample and for non urban residence. Maximum difference of $500 in predicted earnings	557[a]	Closest match within strata detined by age and race. Penalties imposed for previous selection and for non-urban residence. Maximum difference of $500 in predicted earnings	7866

[a]These cases were matched with replacement; 269 different cases were matched a total of 557 times.

[b]These cases were matched with replacement; 336 different cases were matched a total of 786 times.

cases with valid Social Security Numbers and reasonable matches on identifying data were considered.

The random samples were the simplest to construct, as we simply identified CPS cases meeting the specified eligibility criteria. One variation was based upon more comprehensive eligibility criteria than was the other.

In constructing the cell match comparison samples, it was necessary to both determine what the most important cell identifiers were and then to define procedures for collapsing across cells when necessary to attain a critical minimum cell size. This was done on the basis of the order in which measures of various characteristics entered a stepwise regression of post-program earnings estimated on the CPS random samples of young school dropouts and AFDC recipients with no young children. Two variants of the cell-matching procedure were implemented, distinguished by the length of the list of variables considered as potential cell identifiers.

In constructing the statistical match comparison groups, the main issue was defining the appropriate scoring algorithm. In previous applications, two contrasting approaches have predominated: one where the scoring algorithm emphasizes a closeness-of-fit on the predicted outcomes and one where it emphasizes a closeness-of-fit on characteristics correlated with the outcome of interest. Our initial approach matched on predicted earnings measures, where these measures were calculated using coefficients from post-program earnings equations estimated on the CPS samples of young school dropouts and of AFDC recipients with no young children. In our alternative approach, matching was conducted within cells defined by key sample characteristics.

Assessment of the Comparison Groups

In our assessment of the various comparison group methodologies, we focused primarily on the quality of the resulting net impact estimates. Several conditions must be met for a comparison group to yield reliable impact estimates. First, the comparison group must be either similar to program participants along all dimensions that affect the targeted or intended program outcome or different only in ways that have been measured. Second, the program-participant and comparison-group members must exhibit a sufficient degree of similarity so that a common analytic model can be applied to both groups. Third, the comparison group must be large enough to generate reasonably precise estimates of the outcome measures. In constructing comparison groups for the Supported Work Youth and AFDC samples from the CPS, we found potential problems with each of these conditions and, consequently, we found evidence of bias in the net impact estimates.

Not surprisingly, the less narrowly defined random samples of youth and AFDC recipients from the CPS (Stage 1 samples) do not match the program-participant samples well along most measurable characteristics. Furthermore, although the comparability of the participant and comparison

samples was substantially improved by limiting the comparison samples to CPS observations that met the more comprehensive Supported Work eligibility criteria for youth and AFDC recipients (Stage 2 random samples and the initial statistical match samples), large differences persisted. Most notably, minorities are greatly underrepresented in both the youth and AFDC random comparison and initial statistical match groups, undoubtedly due to the urban focus of the Supported Work program. Among the youth comparison groups, both females and married youth are overrepresented, since Supported Work attracted predominantly unmarried minority males. The CPS random sample of AFDC recipients with no young children differs from the Supported Work AFDC sample with respect to its higher average age, higher levels of educational achievement, and higher rate of employment in the base year. Only the statistical match samples obtained by matching on predicted earnings within cells were reasonably comparable in their characteristics to the program participant groups. However, because of the overall sample-size limitations of the CPS, this better match arises mainly from the repeated matching on a subset of the CPS cases included in the other comparison samples.

We also observed a striking differences between the earnings profiles of the youth control and comparison groups, with that for the comparison groups being consistenly much steeper. The earnings profiles for the AFDC control and comparison groups are much more similar to one another.

In view of the pervasive differences in the characteristics of the participant and comparison samples and the observed differences in the earnings trends over time, especially for youth, it is not surprising that tests of the comparability of the earnings models applicable to the comparison samples and to the control group were failed nearly half of the time. The results of these tests suggest that, once the basic demographic and background characteristics are controlled for, some of the comparison groups are sufficiently similar to the experimental group that there is no objective basis for ruling out conducting the net-impact analysis, while other groups are not sufficiently similar.

As seen in Table 2, when these various comparison samples are the basis for estimating net program impacts, there is substantial variability across the estimates. Using results based upon the Supported Work control group as a reference, for youth there is strong evidence of large negative biases in the net impact estimates based on the comparison samples. For AFDC recipients, the qualitative results and often the point estimates of the net impacts obtained using the comparison groups more clearly mirror those obtained using the control group. These basic results held up under both the introduction of other variants of the comparison group construction procedures and changes in the analytic model.

TABLE 2

SUMMARY OF NET IMPACT ESTIMATES GENERATED
USING ALTERNATIVE COMPARISON SAMPLES

	Supported Work Control Group	Random Samples		Cell Match Samples		Statistical Match Model	
		Stage 1 Screens	Stage 2 Screens	Short Match List	Long Match List	Initial	Stratified
Youth							
1977 Earnings[a]	313*	166	-624*	-388	-668*	-511	-774
1978 Earnings	-28	-789	-28	-699*	-1,191**	-408	-873
1979 Earnings	-18	-1,937**	-890	-687*	-1,179**	-339	-1,146
AFDC Recipients							
1977 Earnings[a]	1,423**	1,568**	1,543**	1,696**	1,560**	1,266**	1,256**
1978 Earnings	505**	872**	472	782**	537	515	345
1979 Earnings	351	710	584	806**	257	683	723*

NOTE: These results are reproduced from Table IV.3, IV.4, and E.10. All are based on a basic earnings model, estimated using weighted least squares techniques.

[a]This year, in particular, includes much in-program earnings.

*Statistically signficant at the 5 percent level.

**Statistically significant at the 1 percent level.

Implications for Future Research

The general conclusion from this study is that one should use ex-post comparison group methodologies with extreme caution, particularly in evaluations of youth programs. Although the Supported Work samples differ somewhat from the CETA or JTPA populations, there is substantial overlap in the characteristics of the client populations served by these programs, with the Supported Work samples tending to be representative of the most severely disadvantaged clients in these other programs. For this reason, the CPS is also expected to be an inadequate source of comparison samples for at least certain segments of the CETA and JTPA populations. Lending support to this argument is the fact that the ranges of estimates of CETA net impacts for youth and for women reported in the literature and in companion studies to this are qualitatively, and often quantitatively, similar to those we estimated for the Supported Work Youth and AFDC samples using our comparison groups: CETA net impact estimates tend to be large and negative, for youth (-122 to -1303 dollars per year), and large and positive for AFDC recipients (246 to 949 dollars per year). In view of both the consistency in the patterns of negative estimated net impacts for Supported Work and CETA youth generated from these independently constructed comparison samples and the implausibility that either CETA or Supported Work really had large negative impacts on youth, it seems reasonable to conclude that the basic results from this study for youth may also generalize to youth who enroll in CETA or JTPA.

For AFDC recipients, it appears that any of the comparison group construction procedures will yield reliable indicators of the sign of net impacts. However, the size of the point estimates will be sensitive to the comparison group construction methodology.

Two actions could be taken to improve the future prospects for using comparison group methods in evaluating employment and training programs. The first is to develop a data base that includes a larger number of disadvantaged workers. The second is to ensure that basic employment program eligibility criteria and participation data are measured in the survey. These actions could most efficiently be accomplished through sample and data supplements to recurring surveys such as the CPS or the Survey of Income and Program Participation. However, clearly the safest evaluation strategy involves the use of a true control group since, even with larger samples and better screening data, important unmeasured differences between the program participants and nonparticipants may persist and lead to nonrepresentative comparison samples.

ACKNOWLEDGEMENTS

Many people, both outside of and within Mathematica Policy Research, Inc., deserve recognition for thier contributions to this research. We are especially grateful to our Project Office, Dan Ryan, for this support throughout this project and for his sharing with us the results of his related research. Marilyn Rowan of Westat, Inc. patiently received and answered our many questions about the CLMS and CPS data files. Richard Wehrly of the Social Security Administration skillfully processed our request for Social Security earnings data on the Supported Work Sample. Terry Johnson of SRI, International generously shared with us SRI's preliminary research findings from their companion project for the Department of Labor and their updated CPS data files. In addition, both Terry and Vince Geraci from the University of Texas maintained an open dialogue with us throughout the contract period regarding research strategies, problems, and results and they provided very useful comments on a draft of this report. We are also grateful to Mr. Seymour Brandwein of the National Academy of Sciences and to Harold Beebout and John Burghardt of Mathematica Policy Research, Inc. for their thoughtful reviews of a draft of this report.

Marilyn Manser, currently at the Bureau of Labor Statistics, was the Project Director and a Co-Principal Investigator on this project during the study design period, and, thus, deserves recognition for her role in shaping the project. Robert Moffitt of Brown University and Mathematica Policy Research, Inc. has provided ongoing technical assistance and review of project work. Finally, and most importantly, Beth Fineman and, especially, Douglas Birdzell deserve special recognition for their expert programming and research assistance.

I. INTRODUCTION AND SUMMARY

In recent years, numerous efforts have been devoted to assessing the effectiveness of various employment-intervention strategies. However, because these studies have involved a variety of analytic strategies and data sources, the reliability of their net program-impact estimates has differed greatly. Unfortunately, the most rigorous of the evaluation designs are those associated with a small number of demonstrations that adopted a true experimental design--for example, the National Supported Work demonstration (Hollister et al., 1984), the Louisville WIN Labs project (Goldman, 1981), and the Michigan Private Employment Agency demonstration (Ohls and Carcagno, 1978). Virtually all of the evaluations of CETA and its precursor programs have been based on analytic methodologies that use no comparison group at all or that have defined comparison groups ex post from existing sampling frames (see, for example, Ashenfelter, 1979; Westat, 1980; King and Geraci, 1983; Bassi, 1983; Bassi et al., 1983; Dickinson et al., 1984; Kiefer, 1979; and Mallar, et al., 1982). These non-experimental-design studies suffer from one major limitation: the integrity of their results rests on untestable assumptions about the adequacy of the analytic model and the unmeasurable characteristics of the participant and comparison groups. Nonetheless, as the best estimates available, the findings from these evaluations have been influential in directing public-employment policy.

The purpose of this study is to provide an empirical assessment of the reliability of net program-impact estimates generated through this variety of non-experimental methodologies, and to provide some guidance about which criterion measures should be used to judge the quality of the

analytic results. In the remainder of this chapter, we briefly discuss the program evaluation strategies commonly used to assess employment-related programs or interventions, provide an overview of the goals of, the analytic approach for, and the findings from this study, and describe the organization of the remainder of the report.

A. PROGRAM EVALUATION STRATEGIES

The basic approaches for conducting impact evaluations can be categorized as using (1) a control group, selected within the context of a true experimental design, (2) advance match comparison groups, (3) ex-post comparison-group construction, and (4) internal comparison groups (before-and-after comparisons).

1. True Control Group

Analytically, an experimental design which generates randomly selected participant and control groups offers the most certain opportunity for obtaining unbiased estimates of true program impacts. With such a design, simple differences in means will yield reliable estimates of program impacts.[1] In such cases, the use of statistical modeling is necessary only to improve the efficiency of the impact estimates and to facilitate estimating differential program impacts among participant subgroups.[2] The primary limitation of this procedure is its impracticality

---

[1] If the probability of assignment to the experimental group were other than .50, it would be necessary in constructing the mean values to weight each observation by the probability of cell assignment.

[2] Through the use of statistical modeling, it may also be possible to estimate impacts for synthetic sample subgroups (i.e., subgroups defined by combinations of characteristics not observed with sufficient frequency to allow meaningful comparisons to be drawn between participants and controls who exhibit those particular characteristics.)

in evaluating a fully implemented nationwide program, as was the case for CETA and is currently the case for both JTPA and WIN. For these programs, it is neither feasible nor desirable to deny participation for the purpose of generating a control group.

2. Advance Match Comparison Groups

Probably the next best strategy to using a true control group to estimate net program impacts is the pre-selected, matched comparison-group approach. This procedure, which has been used in, for example, the EOPP demonstration (Brown et al., 1983), the Youth Incentive Entitlement Pilot Projects (Farkas et al., 1982) and the Job Corps Evaluation (Mallar and Kerachsky, 1978), entails identifying a sample of current or former program participants for study and then identifying a group of individuals who are similar in important respects (for example, age, economic background, economic and/or school status) but for whom external barriers to program participation exist (most notably those pertaining to geographic location). Primary data collection (generally longitudinal) is then conducted for both the participant and comparison groups over a comparable time period.

As compared with the true control group, this procedure is less likely to generate a comparison group that accurately reflects what the behavior of and outcomes for participants would have been had they not enrolled in the program. However, it does represent a next-best strategy in cases where a clearly identifiable factor significantly affects the likelihood of participation but is uncorrelated with the measures of program success. In such cases, a variety of statistical modeling approaches will likely generate reasonably reliable estimates of program

impacts. Among the most commonly used procedures are (1) estimating standard earnings models (auto-regressive or cross-sectional) that include a participant status variable, using either regression or maximum likelihood estimation methods, or (2) estimating earning models that incorporate one of the standard selection-bias correction procedures.[1]

3. Ex-Post Matched Comparison Group

The most common approach for estimating net program impacts has involved selecting a comparison sample for the participant group of interest from an existing data base.[2] In developing such comparison samples, research teams have used three general types of procedures. The first is simply to impose some general "eligibility" screens to the data set from which the comparison group is being drawn. The second is to use a cell-match procedure. The third is to use one of two statistical matching techniques--matching along predicted values of participant status or outcome variables, or minimizing a distance function defined as a weighted sum of the differences in characteristics between participants and comparison observations.

---

[1] See, for example, Mallar et al. (1982), Heckman (1979), Barnow et al. (1978), and Maddala and Lee (1976).

[2] For example, Ashenfelter (1979) drew on the Continuous Work History Sample (CWHS) to define a comparison sample for MDTA participants; Cooley et al. (1979) developed a comparison group from MDTA "no-shows"; Gay and Boris (1980) and Kiefer (1979) used a comparison group for various groups of DOL-program participants developed by OEO and DOL from a special survey of residents in selected SMSAs; and Westat (1980), King and Geraci (1983), Bassi et al. (1983), and Dickinson et al. (1984) have all relied on comparison groups for the CLMS sample of CETA participants drawn from CPS data.

Having drawn comparison samples in one of these manners, analysts have then used a variety of analytic approaches for estimating program impacts. The simplest approach has been to compare earnings gains between the pre- and post-program periods. Other approaches have included estimating standard auto-regressive earnings models, random-effects models, and fixed-effects models with and without corrections for serial correlation. In most cases, these models have included actual measures of participation and no selection-bias correction terms. However, in selected instances, statistical procedures for controlling for selection bias were also incorporated into the models.

4. Internal Comparison Group

Few recent program evaluations, but a number of previous ones (see, for example, Perry et al., 1976), have relied solely on the participant sample to estimate net program impacts. In such cases, the analytic approach has been either to estimate pre-post differences in the outcome measures or to estimate shifts in time trends in outcome measures between the pre- and post-program periods. Of all of the approaches, the internal comparison group poses the greatest analytic challenge because the effects of calendar time, aging, and experience gains may all be confounded with the program effects.

B. OVERVIEW OF THIS STUDY

The objective of this study is to assess the quality of the ex-post comparison groups constructed on the basis of various techniques--eligibility screens, cell matching, and statistical matching--that have been used previously or are currently being used in companion projects

funded by the Department of Labor to develop new estimates of CETA net impacts.

Our approach for conducting such an assessment involves two main tasks. The first is to replicate for both the AFDC and youth experimental groups from the Supported Work demonstration a number of the alternative comparison-group-construction procedures that have been used or are proposed for use with the CLMS data. The second is to conduct statistical tests of the differences in evaluation results (particularly net program-impact estimates) based both on the true control groups for these Supported Work target groups and on the various comparison groups. Similarities between the CLMS and Supported Work samples along some key dimensions, the comparability of the data available to conduct the study (including the availability of SSA earnings data up through 1979), and the overlap in the periods of program operations suggest that information on the reliability of Supported Work impact estimates obtained using ex-post comparison groups may also apply to analogous estimates of CETA impacts, especially those for out-of-school youth and welfare dependents. Below, we briefly outline the comparison-group methodologies evaluated, elements of our empirical assessment of the quality of the comparison groups, and finally the main conclusions from the study.

1. The Comparison Group Construction Procedures

We selected the particular comparison-group-construction procedures to be evaluated by considering three factors: (1) the frequency with which various procedures are used in evaluations; (2) existing evidence about the comparability of alternative procedures in terms of the resulting net impact estimates; and (3) the practicality of implementing various

procedures. In total, we selected three types of procedures to be tested for each sample--(1) random sampling of CPS cases; (2) stratified cell-matching; and (3) statistical matching--and incorporated variants into each of the procedures. The variants in the random sample procedures included different screens (less versus more restrictive), those in the cell match sample procedures related to different sets of stratification variables (short list versus long list) and those for the statistical matching included one-on-one matching with no stratification and no replacement versus matching within sample strata with replacement (see Table I.1).

The particular procedure for which numerous other variants could have been examined pertain to statistical matching. Among those variants are (1) matching on the basis of predicted program-participation measures (see, for example, Ryan, 1983) and (2) using one of the traditional distance-function models (see, for example, Alter, 1974; Armington and Odle, 1975; Okner, 1972; Paas, 1982; Kadane, 1978; Rubin, 1979; and Dickinson et al., 1984). However, we chose not to examine the first of these variants because, due to the small-scale demonstration nature of the project, it would have been very difficult to define a reasonable model and sample of nonparticipants to use in predicting Supported Work participation. For the most part, participants in that program were recruited or referred, rather than having responded independently to an ongoing program. Our primary reason for not pursuing the weighted distance functions is the evidence provided by Dickinson et al. (1984) that this substantially more complex and expensive procedure would yield comparison

TABLE 1.1

SUMMARY OF KEY FEATURES OF COMPARISON GROUP CONSTRUCTION PROCEDURES EVALUATED

Procedure/Variant	Supported Work Target Group: Youth	Supported Work Target Group: AFDC Recipients
1. Random Samples of CPS Cases		
Stage 1 screen sample	Restricted to youth who are ages 16-20	Restricted to female AFDC recipients
Stage 2 screen sample	Restricted to youth who are ages 16-20 and are school dropouts	Restricted to female AFDC recipients with no child younger than age 6
2. Cell Match Samples from the CPS		
a. General features	Restricted to youth who are ages 16-20 and are school dropouts	Restricted to female AFDC recipients with no ch younger than age 6
	CPS and Supported Work cases were assigned to cells of 4 or more cases possessing a unique combination of characteristics. Where necessary to achieve cells of 4 or more cases, cel were collapsed across characteristics in the inverse order in which the variables entered a step-wise earnings regression.	
	Cell weights (WGT) computed as $\frac{\#\ SW\ Cases_i}{\#\ SW\ Cases} \Big/ \frac{\#\ CPS\ Cases_i}{\#\ CPS\ Cases}$, where i denotes the cell.	
b. Variants		
Short match list sample	Cells defined by (1) sex, (2) race, and (3) years of school completed.	Cells defined by (1) years of school completed and (2) race.
Long match list sample	Cells defined by (1) sex, (2) earnings 2 years prior to sample enrollment, (3) earnings in base year minus earnings 2 years prior to enrollment, (4) race, (5) years of education, and (6) age.	Cells defined by (1) base year earnings minus earnings 2 years prior to enrollment, (2) age, (3) employed in base year, (4) earnings 2 years prior to enrollment, and (5) race.
3. Statistical Match Samples from the CPS		
a. General features	Restricted to youth who are ages 16-20 and are school dropouts	Restricted to female AFDC recipients with no young children
	Construction procedures included four steps: (1) estimating an earnings equation for the Supported Work experimental group, (2) using the coefficients on the earnings equation to construct predicted annual-earnings values for both the CPS and Supported Work samples, (3) sorting both samples by these predicted values, and (4) sequentially matching each Supported Work case with the CPS case having the closest predicted value within plus or minu $500 of the value of Supported Work case.	
b. Statistical match equation	Predictors included:[a] o Interview Year o Ages 16-17 o Black o Hispanic o Any work base year Weeks worked base year Weeks looked for work base year Hours worked base year o Occupation-Sales/Clerical o Occupation-Service o Occupation-Professional Earnings Base Year Earnings Base Year-1 Earnings Base Year-2 o Per capita Family Income-Medium o Per capita Family Income-High o Less than 10 years of school o AFDC Recipient Base Year o Married o Male	Predictors included:[a] o Interview Year Age Age Squared o Black o Hispanic o Any work base year Weeks worked base year Hours worked base year o Occupation-Sales/Clerical o Occupation-Service o Occupation-Professional Earnings Base Year Earnings Base Year-1 Earnings Base Year-2 o Per capita Family Income-Medium o Per capita Family Income-High o Less than 12 years of school o 12 years of school o No dependent younger than 12

TABLE 1.1 (Continuation)

Procedure/Variant	Supported Work Target Group	
	Youth	AFDC Recipients
c. Variants		
One-on-one match	CPS cases could be selected only once.	
Matching with replacement and within cells	Matching within cells defined by sex and race/ethnicity. Penalties imposed for previous selection and for non-urban residence.	Matching within cells defined by age and race/ethnicity. Penalties imposed for previous selection and for non-urban residence.

o denotes a binary variable

[a] In the variants that matched within sample strata, the sample stratification variables were omitted from the statistical match equations.

groups that would be essentially equal in quality to those constructed on the basis of rather straightforward cell-matching techniques.[1]

A priori, each of these three general comparison-group-construction procedures has strengths and limitations relative to the other. For instance, the random sample approach can easily be implemented and yields the largest samples. However, it imposes the greatest burden on the analytic methodology used to estimate net impacts; virtually all analytic results must be generated through statistical modeling. The second approach (cell matching) is relatively straightforward to implement, and, by weighting the comparison-group observations, analysts can find some value in the descriptive statistics for the participant and comparison samples. However, unless one assumes that the analytic model fully controls for the unique features of each sample cell, either weighted statistical analysis must be undertaken or cell assignments must be controlled for directly. The third appproach, statistical matching, is difficult to implement and, in one of the variants we have explored, does not ensure that sample members will be well matched along characteristics. Finally, these and other common variants of statistical matching also assume that the earnings model from which scoring weights are generated applies to both the participant sample and the entire sample from which the comparison group is being drawn. This assumption may be tenuous. However, this matching approach is advantageous because the one-on-one matching of

---

[1]In addition, the results of other evaluations of statistical matching procedures point out that the sensitivity of the match quality depends critically on the functional form of the distance function but not on the complexity of the scoring algorithm (see, respectively, Rodner and De Vol, 1982, and Barr et al., 1981). For this reason, we viewed the test of alternative functional forms as the more interesting variant.

sample members on the basis of predicted outcomes or the closeness of their characteristics eliminates the necessity of conducting weighted analysis.

2. The Evaluation of Alternative Comparison-Group Methods

The quality of the various comparison groups constructed for the Supported Work youth and AFDC samples can be assessed along several dimensions. How do the comparison-group members compare with the participant-group members in terms of their background characteristics, employment history, and employment behavior? How similar are the time trends in earnings between the comparison-group members and the true control-group members? How do estimates of net program impacts that are generated by using the participants and the true control group compare with those that are estimated by using the comparison groups? Are the net-impact estimates sensitive to the specification of the estimation models? Of particular importance in our ability to address these questions is the availability of a control group for each of the Supported Work samples to provide the necessary benchmarks against which to compare the comparison group in all time periods when the behavior of and the outcomes for the participant group will reflect, in part, any influence of the program.

Our approach for addressing each of these questions and the results of the analysis are summarized descriptively in Table I.2. Insofar as has been practical and feasible, we have followed quite closely the models and assumptions common to related studies of CETA net impacts based on the CLMS data. Doing so maximizes our ability to compare our results with these other studies and to generalize our findings so as to offer some assessment about the quality of the net-impact estimates of CETA being generated in these studies.

TABLE I.2

DESCRIPTIVE SUMMARY OF EVALUATION
ISSUES, ANALYTIC APPROACHES, AND HIGHLIGHTS OF RESULTS

Issues	Analytic Approach	General Results: Youth — Random Samples	Youth — Cell Match Samples	Youth — Statistical Match Sample	AFDC Recipients — Random Samples	AFDC Recipients — Cell Match Samples	AFDC Recipients — Statistical Match Sample
1. Comparability of Participant and Comparison-Group Members							
a. Demographic and background characteristics	Descriptive statistics (weighted for the cell match samples)	In all comparison samples, minorities are underrepresented and females and married youth are overrepresented as compared with the participant sample.			In all comparison samples, minorities are underrepresented, as are younger women and women with fewer than 12 years of education as compared with the participants.		
b. Employment behavior	Chow tests of the similarity of estimated earnings models for comparison groups and the control groups	Failed test with Stage I screen sample.	Failed test with short match list sample.	Passed test.	Failed all tests.	Passed all tests.	Passed test.
2. Similarity in the Time Trends in Earnings	Comparison of mean annual earnings for the the control group and for the comparison groups (regression-adjusted and not)	Substantially steeper earnings gains over time and generally much higher levels of earnings in the outcome years among comparison group members as compared with the Supported Work control group.			Generally similar earnings trends over time for the control and comparison groups.		
3. Similarity of Net Impact Estimates							
a. Unadjusted estimates	Comparison of differences in mean earnings gains between the participants and the comparison groups and between the participants and the controls	LN all years	LN all years	LN all years	SP 1977 MP 1978 LP 1979	SP 1977 MP-LP 1978 SN-LP 1979	SP 1977 LP 1978 LP 1979
b. Basic earnings model estimates	Comparison of regression-adjusted differences in mean earnings estimated from the participant and comparison samples as compared with the participant and control-group sample	SN-LN 1977 LN 1978 and 1979	LN 1977 and 1978 SN-LN 1979	LN all years	SP 1977 SN-LP 1978 O-LP 1979	SN-SMP 1977 MP 1978 SN-MP 1979	SN 1977 O 1977 SN-LP 1979
c. Fixed-effects earnings model estimates	Comparison of regression-adjusted differences in means earnings gains estimated from the participant and comparison samples as compared with the participant and control-group sample	LN all years	SN-LN entry year. LN all other years.	SN-MN earlier	SP-LP most recent years. A SN early years.	SN-SP early years. MP-LP later years.	
4. Sensitivity of Net Impact Estimates to Alternative Assumptions							
a. Applicability of different analytic models to various subsets of the population	Comparison of results from basic earnings model estimated for sample subgroups and comparison of results from the Basic Earnings Model with various interaction variables added	Similar conclusions on quality of matches as with overall sample.					
b. Effects of Weighting for Grouping	Comparison of estimates on partially versus fully grouped data.	*Minimal effects for grouping data.*					

TABLE 1.2 (Continuation)

Issues	Analytic Approach	General Results					
		Youth			AFDC Recipients		
		Random Samples	Cell Match Samples	Statistical Match Sample	Random Samples	Cell Match Samples	Statistical Match Sample
c. Sensitivity of cell match results to the weights used	Comparison of results estimated on weighted and unweighted data	Minimal effects for using cell match, group weights, and for combined weights.					

SP means small positive bias in impact estimates.
MP means medium positive bias in impact estimates.
LP means large positive bias in impact estimates.
SN means small negative bias in impact estimates.
LN means large negative bias in impact estimates.

3. The Major Lessons and Conclusions From this Study

The primary lessons from this study can be grouped into those that pertain to (1) the adequacy of the CPS for constructing comparison groups, (2) the relative desirability of alternative comparison-group-construction methods, (3) the sensitivity of net impact estimates to the analytic model specification, and (4) the overall reliability of net-impact estimates generated with comparison groups.

The Adequacy of the CPS. Several features of the CPS may seriously limit its utility as a source of comparison samples for the Supported Work samples and for other groups of employment program participants. These limitations relate to the size of the CPS sample, its geographic distribution, and the unavailability of some important data items.

Although the CPS is a very large survey, it includes relatively few observations on individuals in some population groups that have a high probability of participating in employment programs. Each Current Population Survey contains over 7,000 youth, but only about 1,200 young school dropouts. Similarly, fewer than 1,000 female welfare recipients are contained in each. Furthermore, a large portion of the disadvantaged CPS observations reside outside of the central cities where many of the special employment initiatives are undertaken. Consequently, the CPS contains relatively few cases that match out-of-school youth and welfare recipients who enroll in employment programs along a combination of key characteristics. This sample may better serve as a basis for constructing comparison samples for school youth and for adult males. However, the number of CPS cases within even these groups who are similar to program

participants along dimensions relevant to their employment success may also be quite limited.

Perhaps the most important data elements that are missing from the CPS are those that relate to participation in either the employment program under study or in complementary or alternative programs. The result of this is some degree of contamination of the comparison groups.[1] This can be an especially serious problem when studying youth programs operated during the mid to late 1970s, when a large portion of disadvantaged youth participated in some type of employment program. On the one hand, failure to measure program participation among comparison group members will tend to result in downwardly biased estimates of the true program impacts. On the other hand, in the case of evaluating national programs like CETA, to the extent that the presence of these employment intervention programs may have had sizeable impacts on the private sector employment opportunities available to nonparticipant youth, the net impact estimates may understate true program impacts.[2]

Other important data limitations of the CPS include the unavailability of geographic identifiers, local economic indicators, and earnings data for briefer time periods than a year. Each of these factors potentially would be an important sample selection or matching criteria.

---

[1]Bassi et al. (1984) estimate that this type of contamination may bias net CETA impact estimates substantially since 32 to 50 percent of women and 35 to 68 percent of youth who are selected for inclusion in the comparison samples will have previously participated in an employment program.

[2]That is to say that the withdrawal of large numbers of youth from the private-sector work force may have opened positions for the remaining youth.

Alternative Comparison Group Construction Methods. When multivariate statistical techniques are used to obtain estimates of net program impacts, the actual comparison-group-construction method appears to be less important than is the definition of the subset of cases eligible for inclusion in the comparison sample. Both in our study and in a companion study conducted by SRI International, Inc. (Dickinson et al., 1984), strong evidence suggests that the general sample screens are important, but that using a well-screened subset of the general population tends to yield net-impact estimates that are not qualitatively worse than estimates based on statistical-match or cell-match comparison samples drawn from these screened samples. However, when net program impacts are estimated simply by examining differences in mean earnings between participant and comparison groups, matched comparison samples yield more reliable results than do screened comparison samples.

Analytic Model Specifications. The net-impact estimates are sensitive to major changes in the analytic models. For example, basic earnings models that control for background characteristics and earnings histories tend to generate better estimates than do models that use only program participation to explain the change in earnings between a pre-program year and the outcome year. Furthermore, the basic earnings models tend to yield less biased estimates than do fixed-effects models. However, minor changes in the basic earnings models to include additional variables or to alter the definitions of variables have minimal effects on the net-impact estimates.

Overall Reliability of Results Based on Comparison Group Methodologies. The results of this study indicate that net-impact estimates may differ substantially between those generated using control groups and those generated using comparison groups, especially for the young school drop-out sample. Furthermore, the comparison-group results sometimes vary substantially across comparison groups and analytic models. For youth, this instability appears to be attributable to the fact that those who participated in Supported Work had substantially different age-earnings profiles than are typical of the entire population of young school dropouts. These differences apparently were not captured by either our cell matching or our statistical matching criteria and, perhaps, are simply not measurable attributes. For AFDC recipients, the earnings trends are reasonably similar between the control and comparison groups. However, the point estimates of net impacts still vary some across comparison groups.

In relating the results of this study to the estimates of program impacts generated in previous research efforts, we conclude that, at least for youth, one must be skeptical of the reliability of estimates based on comparison group methods. For welfare recipients, we can be somewhat more optimistic that the qualitative conclusions from such estimates will be valid.

C. ORGANIZATION OF THE REMAINDER OF THE REPORT

The remainder of this report consists of three chapters. Chapter II describes the sample and the data. Chapter III outlines the comparison-group-construction procedures and summarizes the results of each sample-construction effort. Finally, Chapter IV presents the results of our analytic assessment of these various comparison groups constructed for the

Supported Work youth and AFDC samples, and discusses the implications of these findings for interpreting the results of previous net impact studies and for evaluation-design strategies aimed at other programs, such as those funded under JTPA.

II. SAMPLE AND DATA

The success of a study to assess alternative comparison-group methodologies for evaluating CETA-type programs depends critically on the data set. The data should allow true program impacts to be estimated, and it should include a program participant group that is similar in important respects to participants in previous CETA and current JTPA programs. Furthermore, it is important that comparable data bases be available for the program participant sample and the sample from which comparison groups are to be drawn. The national Supported Work data base is especially appropriate in the first two respects, while the Current Population Surveys (CPS) provide a reasonable data set from which comparison groups can be developed.

Our methodological work has focused on developing comparison groups from the CPS for the youth and AFDC participant groups in the Supported Work samples. In Section A, we describe the Supported Work youth and AFDC samples and the Supported Work data base used in the study. In Section B, we describe the CPS samples used. Finally, in Section C, we outline our general approach to combining the CPS and Supported Work data sets.

A. THE SUPPORTED WORK YOUTH AND AFDC SAMPLES

We chose the Supported Work youth and AFDC samples for four reasons. First, these samples consist of individuals who tended to have serious labor-market disabilities and who were characterized as economically disadvantaged, and, thus, who tended to possess characteristics that were typical of the target population for both CETA and JTPA. Second, the Supported Work program enrolled its participants

between April 1975 and July 1977, a period of time which is consistent with both the period of CETA operations covered by CLMS data and the CPS data files constructed for use with the CLMS data. Third, the Supported Work sample includes a randomly selected control group, which will enable us to generate unbiased estimates of the true program impacts for the Supported Work youth and AFDC groups. These unbiased estimates of impacts can then be used as benchmarks for judging the quality of the estimates obtained from comparison groups constructed using the alternative methodologies. Finally, the Supported Work and CPS data bases are reasonably comparable, and both data sets have been augmented with Social Security earnings data.

1. The Supported Work Demonstration and Sample Design

The national Supported Work demonstration, conducted between 1975 and 1979, was a large-scale field test of the effectiveness of a highly structured work-experience program in mitigating the employment problems of persons with severe employment disabilities, including young school dropouts and AFDC recipients. The Supported Work employment experience differed from other programs targeted toward these groups in one major respect: it emphasized the provision of up to 12 to 18 months of work experience that consisted of peer-group support, graduated stress, and close supervision. Integral to this demonstration was a rigorous research effort that included an experimental design whereby, between April 1975 and July 1977, eligible program applicants in 10 sites were randomly assigned to one of two statuses: program participant or control-group member. Individuals in both groups were then interviewed at enrollment and at

subsequent 9-month intervals through March of 1979, yielding 18 to 36 months of follow-up data.[1]

The final data base constructed for the research sample included (1) data on demographic characteristics, employment history, welfare dependence, and household composition as of the time of enrollment in the sample;[2] (2) data on employment and welfare dependence, derived from the follow-up surveys; and (3) Social Security earnings data for the years 1972 through 1977. As part of the current study, Social Security earnings data for 1978 through 1981 were added to the data base. An important feature of the data base is that the Social Security data provide mean values for cells of 7 to 10 sample members, where sample members were assigned to cells on the basis of (1) target group (AFDC, youth, ex-addict, or ex-offender), (2) experimental status, (3) date of enrollment (prior to or not prior to April 1976), (4) whether or not they were employed in the year prior to enrollment, and (5) site.[3]

---

[1] Those enrolled in the sample prior to April 1976 were to have generated 36 months of follow-up data; those enrolled between April and December 1976 were to have generated 27 months of follow-up data; and those enrolled during 1977 were to have generated only 18 months of follow-up data.

[2] Data on drug use and criminal histories were also collected for all sample members except those in the AFDC target group.

[3] This cell grouping was based on the prior judgments that neither members of different target groups nor experimentals and controls should be combined in the same cell. Other cell identifiers were selected on the basis of the results of an AID III analysis of earnings data.

However, one should note that (1) all Supported Work youth are school dropouts, while 67 percent of CETA youth were still enrolled in school, and that (2) a much higher proportion of the Supported Work youth had had a "brush with the law." (See also Appendix Table A.1.) In assessing the generalizability of the results from this methodological study to CETA or JTPA youth, these differences between Supported Work and CETA youth should be kept in mind. In particular, the results of this study should be viewed as being most generalizeable to out-of-school CETA and JTPA youth.

3. The AFDC Sample

Seven of the demonstration sites enrolled AFDC recipients: Atlanta, Chicago, Hartford, Newark, New York City, Oakland, and a rural area of Wisconsin. The first eligible AFDC recipient was randomly assigned to a participant status in Newark in February 1976, with random assignment continuing through July 1977, by which time 1,620 women had been enrolled in the AFDC sample (804 experimentals and 816 controls).

Eligibility for the AFDC target group was limited to females who had a long history of AFDC receipt, limited or no recent employment, and no young child (see Table II.5). As a result of these criteria and the program locations, the subsequent sample was drawn primarily from large cities and consisted of women who exhibited numerous characteristics that are associated with employment problems. As indicated in Table II.6, the typical person in the sample was 34 years old, was black, had 10 years of schooling, had two dependents, and had a youngest child between the ages of 6 and 12. The average welfare payment was about $280 per month, with an additional average food stamp bonus value of about $70. The women had received welfare for almost nine years on average, and the average length

TABLE II.5

ELIGIBILITY FOR ENROLLMENT IN THE SUPPORTED WORK AFDC TARGET GROUP

	Restriction
Sex	Female
Recent Employment	No more than 40 hours worked during the last four weeks
Employment History	Spent no more than three months in one regular job (a job of 20 or more hours per week) during the last six months
	Did not previously participate in either a Supported Work program or the control group for the evaluation component of the Supported Work demonstration
AFDC Receipt	On AFDC continuously during past three years--that is, on AFDC (1) both at present and 36 months ago, and (2) for at least 30 of the past 36 months
Age of Children	Youngest child at least 6 years old

SOURCE: MDRC (1978), Table III.2.

TABLE 11.6

CHARACTERISTICS OF THE SUPPORTED WORK AFDC SAMPLE USED IN THE METHODOLOGICAL ANALYSIS

	Experimentals	Controls
Age		
Percent younger than 21	1.00	0.50
Percent 21-34	58.13	59.98
Percent 35 or older	40.75	40.02
(Average)	(33.33)	(33.46)
Education		
Percent with fewer than 12 years	69.70	69.20
Percent with 12 or more years	30.30	30.80
(Average years)	(10.28)	(10.27)
Race/Ethnicity		
Percent black, non-Hispanic	83.50	81.55
Percent Hispanic	12.00	12.97
Percent other	4.50	5.49
Age of Youngest Child		
Percent younger than 6	2.12	2.62
Percent 6-12 years	71.13	72.32
Percent 12 or older	25.62	23.94
Number of Dependents		
Two or fewer	64.00	62.85
More than two	36.00	37.15
Percent Receiving Welfare--Base Year	100.00	100.00
(Average Monthly Family Income--Base Year)	($309.62)	($309.42)
Percent of Time Employed--Base Year		
0	80.25	78.18
1-49	15.12	16.83
50-89	2.87	3.37
≥ 90	1.37	0.75
Percent of Those Employed Who Were Employed in Public-Sector Jobs	n.a.	n.a.
Average Interview Earnings--Base Year		
None	80.63	79.18
$1-1,000	11.13	12.59
$1,001-5,000	17.75	7.73
> $5,000	0.38	0.37
(Average)	(232.49)	(232.94)
Average SSA Earnings--Base Year[a]		
None	10.64	5.39
$1-1,000	79.35	83.96
$1,001-5,000	10.01	10.65
> $5,000	0.00	0.00
(Average)	($458.57)	($463.24)
Average Change in SSA Earnings--Pre-Base to Base Year[a]		
Negative	47.88	56.23
Zero	16.25	12.34
Positive	35.88	31.42
(Average)	($-55.15)	($-153.27)
Number of Cases[b]	800	802

NOTE: Percentages do not always sum to 100 due to missing data.

[a] These data are based on mean values for cells of 7 to 10 sample members.

[b] Eighteen sample members failed to complete the baseline interview and, hence, have not been included in these statistics.

of time since the last regular job for those with some prior work experience was nearly four years (see Masters and Maynard, 1981, pp. 27-30).

As nearly as can be determined from the 1975 AFDC survey data, approximately 17 percent of all AFDC recipients met the Supported Work eligibility criteria. However, as indicated in Table II.7, those who met the eligibility criteria differ substantially from the entire AFDC population, from WIN participants, and from the Supported Work AFDC sample (especially in terms of their racial/ethnic composition).[1] Furthermore, the Supported Work AFDC sample differs from female participants in CETA in important ways--most notably, the Supported Work sample has lower levels of education and a higher representation of minorities (see Appendix Table A.2). Thus, again, the generalizability of the results from this methodological study to either the overall population of AFDC recipients or to female participants in CETA or JTPA is uncertain. However, they probably generalize quite well to female welfare recipients who participate in those programs.

B. THE CPS SAMPLES USED TO DEVELOP COMPARISON GROUPS

The selection of CPS samples as the data source for generating comparison groups for the Supported Work participant groups was based on several considerations. First, the March CPS collects detailed data on current employment status and work experience during the previous year, as

---

[1] This is undoubtedly due to the concentration of the Supported Work sample in large urban areas.

TABLE 11.7

PERCENTAGE DISTRIBUTIONS OF THE AFDC AND WIN PARTICIPANT POPULATIONS AND OF THE SUPPORTED WORK AFDC SAMPLE, BY DEMOGRAPHIC CHARACTERISTIC

	Adults on AFDC, 1975[a]	AFDC Recipients Meeting Supported Work Eligibility Criteria, 1975[a]	WIN Participants 1974[b]	Supported Work AFDC Sample[c]
Sex				
Male	12.1[d]	0.0	30.0	0.0
Female	87.9	100.0	70.0	100.0
Age				
Under 25	31.3	2.1	43.5	8.3
25 and over	68.7	97.9	56.5	91.7
Race				
White	41.4	33.4	51.1	5.0
Minority	58.7	66.6	48.9	95.0
Education				
Less than 12th grade	62.8	74.4	55.2	69.4
12th grade or more	37.2	25.6	44.8	30.6
Family Size				
2 or less	21.9	20.5	27.6	25.7
3-4	43.3	42.7	47.0	47.9
5 or more	34.8	36.8	25.5	26.4
Months on Welfare				
1-6	11.2	0.0	50.3	0.4[e]
7-12	10.9	0.0	9.4	0.3[e]
13-24	15.9	0.0	12.7	1.7[e]
25 or more	62.0	100.0	27.6	97.6
Months Since Last Full-Time Job				
Now working or < 2	17.6	0.0	7.7	3.2[f]
2-6	3.9	4.2	31.9	5.2
7-12	5.8	4.6	12.8	6.6
13-24	6.2	5.7	11.1	11.3
25 or more (or never worked)	66.5	85.5	36.4	73.6
Number of Persons	2,977,768	494,501[g]	n.a.	1,602

[a] These figures are based on data from a survey of 26,617 AFDC recipients conducted by the Department of Health, Education, and Welfare in 1975. An analysis of these data was conducted by the MPR programming staff.

[b] These data are reported in Schiller et al. (1976) and are based on a random sample of WIN registrants in 78 sites drawn from the longitudinal evaluation of the Work Incentive (WIN) Program.

[c] These data are from the Supported Work baseline interview and include both experimental and controls.

[d] If present, males are assumed to be the AFDC recipients.

[e] The fact that some persons in the AFDC sample were on welfare less than 30 months means either that they were technically ineligible to participate in Supported Work or that there were reporting errors.

[f] A full-time job is defined as any job that consisted of 20 or more hours a week and lasted at least two weeks. (This definition is not consistent with that used to determine eligibility.)

[g] This figure is based on the assumption that the 123,945 recipients whose eligibility status could not be determined due to missing data are represented among the eligible and non-eligible subsamples in the same proportions as those whose eligibility status could be determined.

n.a. means not available.

well as basic demographic and background data.[1] Second, the CPS is a large, nationwide population survey, and thus includes observations on individuals who are similar to the Supported Work participants. Third, appended to the CPS files for the years comparable to the Supported Work data (1976-1977) are Social Security earnings data for 1951 through 1979.[2] Finally, the CPS is the most readily available data set for such purposes; indeed, it has been used in most of the previous and ongoing work to estimate CETA net impacts or to develop performance standards from the CLMS data.[3]

As has been done in previous studies that have used these CPS data to construct comparison groups (for example, Westat, 1980), we restricted the universe of eligible CPS cases for inclusion in a comparison group in several ways. For comparability with the Supported Work enrollment period, only the March 1976 and March 1977 CPS samples were used. Furthermore, the sample was limited only to those between the ages of 15 and 64 (inclusive) for whom valid Social Security data were available for the years 1972 through 1979, and for whom there was an acceptable quality match with respect to key identifiers between the CPS case and the Social Security

---

[1] To some extent, the Supported Work interviews were modeled after the CPS surveys, and, thus, the baseline data for the two samples are reasonably similar.

[2] This was done for Westat, Inc., as part of its efforts to develop public-use tapes for the CLMS files (see Westat, July 1981).

[3] One exception is the work by Bassi et al. (1984), which used the ISDP data for some analysis.

record from which the SSA earnings data were obtained.[1] The results of these screens on the overall CPS sample are summarized in Table II.8.

From this total sample of eligible CPS cases, two subsamples were drawn for subsequent comparison-group constructions and analysis. One was a sample of youth, the other a sample of AFDC recipients.

1. The CPS Youth Sample

We selected from the 1976 and 1977 CPS files all youth between the ages of 16 and 20 at the time of the CPS survey. This sample was used subsequently both as one of the comparison groups for the Supported Work youth sample and as the universe file from which other comparison samples were drawn. As indicated in Table II.9, this sample included approximately 14,000 youth. The majority of these youth were younger than age 19 (59 percent), had not completed high school (56 percent), and had relatively low earnings, even though over two-thirds of the youth held some job during the base year. Relatively few (12 percent) of these youth were married. They share with the Supported Work youth sample their relative youth and their low levels of employment and earnings, but, in other respects, they differ substantially (see Table II.3)--most notably, a much higher proportion of the CPS sample was either still enrolled in school or had completed their high school education.

---

[1] The exclusion of cases without valid Social Security numbers may eliminate many youth who have dropped out of school but never held a covered job. Such cases were important subsets of the Supported Work Youth and CETA target groups. Thus, eliminating such cases from the comparison sample could be a source of bias in the results.

TABLE II.8

SCREENS APPLIED TO CPS FILES TO GENERATE SAMPLES ELIGIBLE FOR INCLUSION IN THE COMPARISON SAMPLES

File/Screen	Number of Cases
CPS - GENERAL SCREEN SAMPLE	**116,245**
1. CPS - March 1976 and March 1977	203,709
CPS-1 (March 1976)	81,323[a]
CPS-2 (March 1977)	122,386[a]
2. CPS - Ages 15-64[b]	
Valid Social Security Earnings Data, 1972-1979[c]	116,245
Data on Personal Characteristics Match Between the CPS and SSA Records[d]	
a. Birth year within one year	
b. Sex	
c. Name, within one letter, if male	
d. Race, if female	
e. Month of birth, if female	
CPS-1	59,476
CPS-2	56,769

NOTE: Screens were applied in the order in which they appear in this table.

[a] The 1977 CPS file received from Westat included cases without Social Security numbers, while the 1976 CPS file did not include those cases.

[b] Roughly 85 percent of the CPS sample were between the ages of 15 and 64 (inclusive).

[c] Roughly 60 percent of the CPS cases had Social Security numbers.

[d] Roughly 90 percent of the CPS cases with Social Security data had acceptable agreement with respect to the identifying variables.

TABLE 11.9

CHARACTERISTICS OF THE CPS SAMPLES OF YOUTH AND AFDC RECIPIENTS

	Youth	AFDC Recipients
Age		
Percent younger than 19	58.92	--
Percent 19 or older	41.08	--
Percent 20 or younger	--	9.42
Percent 21-34	--	56.49
Percent 35 or older	--	34.09
(Average)	(18.06)	(31.61)
Education		
Percent with fewer than 10 years	16.08	59.55
Percent with 10-11 years	40.10	
Percent with 12 or more years	43.82	40.45
(Average years)	(11.99)	(11.11)
Race/Ethnicity		
Percent black, non-Hispanic	10.81	37.84
Percent Hispanic	6.39	11.98
Percent other	82.79	50.17
Age of Youngest Child		
Percent younger than 6	--	54.44
Percent 6-12 years	--	30.22
Percent 12 or older	--	15.34
Number of Dependents		
Two or fewer	--	59.85
More than two	--	40.15
Marital Status		
Percent married	12.57	0.00
Percent not married	87.43	100.00
Percent Receiving Welfare--Base Year	2.04	100.00
(Average Monthly Family Income--Base Year)	($345.99)	($554.40)
Percent of Time Employed--Base Year		
0	29.63	61.55
1-49	31.04	16.74
50-89	17.06	10.93
$\geq$ 90	22.26	10.78
Percent of Those Employed Who Were Employed in Public-Sector Jobs	n.a.	6.17
Average Interview Earnings--Base Year		
None	29.64	61.55
$1-1,000	29.64	13.73
$1,001-5,000	31.60	19.55
> $5,000	9.06	5.11
Average SSA Earnings--Base Year		
None	32.87	59.60
$1-1,000	25.48	15.64
$1,001-5,000	33.36	19.85
> $5,000	8.28	4.91
(Average)	($1500.09	($919.19)
Average Change in SSA Earnings--Pre Base to Base Year		
Negative	17.76	25.11
Zero	29.64	49.57
Positive	52.60	25.31
(Average)	($563.28)	($39.78)
Number of Cases	14,084	1,995

NOTE: Percentages do not always sum to 100 due to missing data or rounding.

2. The CPS AFDC Sample

The CPS AFDC sample includes all females in the CPS sample who had one or more dependents younger than age 18 and who received welfare during the calendar year prior to their CPS interview. As indicated in Table II.9, this sample consists of 1,995 women who were relatively young, disproportionately black, and had very low levels of employment and earnings. They share with the Supported Work AFDC sample their relatively young age, the disproportionate representation of minorities, and their limited work experience and low income (see Table II.6). However, they differ most notably in that many of these women have young children, a much higher proportion have completed high school (40 versus 30 percent), and a lower proportion are black (38 versus 82 percent).

C. THE COMBINED SUPPORTED WORK AND CPS DATA SETS

In preparing the Supported Work and CPS data for analysis, it was necessary to identify relevant information that was common to the two data sets and to devise and implement algorithms for constructing variables with similar meanings and interpretations. Table II.10 summarizes the key variables used in the analysis and identifies key assumptions imposed in aligning the data from the two sources. Although, we would like to have been able to include area unemployment rates in our data set, this information is not available for the CPS sample. We did, however, compare the rates for 1974 through 1979 in the Supported Work sites with those for the nation as a whole and found that, with the exception of two of the sites enrolling each AFDC recipients and youth, unemployment rates in the Supported Work sites tended to be around or below the national average over most of this period.

TABLE II.10

SUMMARY OF ANALYSIS VARIABLES

Variable Description	Comments	
	Supported Work Samples	CPS Samples
A. Outcome Measures		
SSA earnings 1977 SSA earnings 1978 SSA earnings 1979	Cell Means	Individual Values, Grouped for Impact Analyses
B. Descriptive and Control Variables		
Program Status	1 if experimental group, 0 otherwise	0 for all observations
Age	As of date of sample enrollment	As of March of interview year
Ethnicity/Race Black, non-Hispanic Hispanic Other		
Any Work Base Year	12 months prior to enrollment	1975 for 1976 CPS 1976 for 1977 CPS
Weeks Worked Base Year		
Weeks Unemployed Base Year for Those with No Employment [a]		
Average Hours Worker Per Week	4 weeks prior to enrollment	Week prior to survey
Occupation Sales/Clerical Service Worker Professional	DOT codes longest job in 2 years prior to enrollment	Census codes current/ most recent job
Interview Earnings in Base Year	Year prior to enrollment	1975 for 1976 CPS 1976 for 1977 CPS
SSA Earnings in Base Year	cell means	Individual values, grouped for Impact Analysis
SSA Earnings Base Year - N (N = 1,2)		
Annual Family Income Earnings of Sample Member plus Welfare Benefits to All Household Members	Welfare income equals annualized value of benefits received in month prior to enrollment	Pertains to calendar year prior to interview
Years of School Completed		
Number of Dependents	Own children younger than 19	Household members younger than 19
Age of Youngest Dependent		
Current School Status	At enrollment	As of March CPS Survey
Marital Status		
Welfare Receipt Excluding Food Stamps	Month prior to enrollment	1975 for 1976 CPS 1976 for 1977 CPS
Cohort 1976/1977	1976 if enrolled prior to October of 1976	
Central City SMSA	All sites except Wisconsin	Based on Census SMSA codes

[a] This definition was to impose consistency with the CPS questionnaire.

As mentioned previously, a major difference between the Supported Work and the CPS data is the fact that individual Social Security earnings data are available for the CPS sample, while only cell mean values are available for the Supported Work sample. Our approach for dealing with this problem has been simply to group (ex-post) the CPS comparison sample members using a procedure similar to that used to group the Supported Work observations, and then to conduct the impact analysis on the basis of grouped data for both the CPS and the Supported Work samples, using weighted least squares estimation techniques (see Appendix F for a further discussion of the weighting procedures used).

III. CONSTRUCTION OF COMPARISON GROUPS

A. INTRODUCTION

Because the CETA program was not conducted as a formal social-science experiment in which qualified program applicants were assigned randomly to experimental and control groups, its effects on employment and earnings cannot be estimated reliably in a straightforward manner. Several methodologies for drawing pseudo-control-groups (comparison groups) from CPS data have been adopted in recent evaluations of the CETA program (see, for example, Westat, 1982; Bassi et al., 1983; and Dickinson et al., 1984). In these studies, the comparison groups have been used in conjunction with participant groups to estimate the effects of CETA. However, the absence of a true CETA control group precludes a full assessment of the reliability of these estimates.

This chapter describes three different methodologies for obtaining comparison groups with which to evaluate the effects of the Supported Work program on earnings. This employment and training program forms the basis for our assessment of these comparison-group methodologies because it provides data on a control group of eligible persons who were randomly excluded from participating in the program. The comparison-group construction methodologies are evaluated in this chapter by comparing the characteristics of individuals in comparison groups with those of individuals in the experimental and true control groups, and by comparing trends over time in mean earnings. In Chapter IV, regression estimates of program impacts on earnings that were obtained using the comparison groups are compared with those obtained using the true control group. Based on

these evaluations, we make recommendations about using these methodologies to obtain comparison groups for CETA and JTPA participants.

Westat (1982) used stratified matching (cell-matching) procedures to obtain CETA comparison groups from CPS files. Each CETA participant in a CLMS sample was located in a cell defined by categories of as many as 12 variables that determine earnings. A contemporaneous CPS data file was searched to locate matching cases. If none was found, the cell was redefined to include additional categories of some variable. Categories of variables having the weakest effects on earnings were the first to be combined. This continued until matching CPS cases were found. The outcome of this process was a comparison group of CPS cases that closely resembled the CETA experimental group with respect to characteristics that were known to affect earnings.

Dickinson et al. (1983) used a statistical matching (nearest-neighbor matching) procedure to draw CETA comparison groups from the CPS. This procedure was based on a Mahalanobis distance function, which is a modified sum of the squared weighted deviations of selected characteristics of a potential comparison case (CPS) from those of a treatment case (CLMS). The selected characteristics were those which, as shown by a regression analysis, were significant in explaining earnings. The associated weights were modified beta weights (standardized regression coefficients) from the earnings regression. Each CLMS case was matched with the CPS case that minimized the value of the distance function. As with stratified matching, the result was a comparison group of CPS cases that closely resembled a CLMS sample of CETA participants with respect to characteristics that determine earnings.

An alternative statistical matching procedure would be to match each treatment case with the comparison case that is most similar with respect to predicted earnings. Doing so would produce a comparison group that resembles the treatment group in predicted earnings, but that may differ with respect to factors that determine earnings.

Stratified matching and the second variant of statistical matching (matching on predicted outcomes) are among the comparison-group-construction methodologies which we describe and evaluate in this and the following chapter. We also evaluate a third and much simpler comparison-group-construction procedure (random samples of youth and AFDC recipients) to determine whether substantial additional benefits are derived from the more sophisticated procedures.

The remainder of this chapter describes these three methodologies for obtaining Supported Work comparison groups from the CPS. We first discuss the random selection of CPS cases that satisfied Supported Work eligibility criteria. We then discuss the implementation of stratified matching. The chapter concludes with an examination of a statistical matching procedure that utilizes results from a regression analysis of earnings.

B. RANDOMLY SELECTED COMPARISON GROUPS

If data are available both for a group of participants in an employment and training program and for a carefully selected comparison group, then crude estimates of program impacts could be obtained by comparing the mean post-program earnings of the two groups. More sophisticated estimates could be obtained by using regression analysis to control

for individual differences in characteristics that affect earnings. The less similar the comparison and participant groups are with respect to these characteristics, the more important it is that the program-impact estimator be incorporated into a fully specified model of earnings.

Our simplest comparison groups for youth and AFDC women who participated in the Supported Work program consist of random samples of CPS cases that met minimum data-quality standards and satisfied certain program-eligibility requirements. For these comparison groups, no attempt was made to match CPS cases with Supported Work cases on the basis of factors that affect earnings. Consequently, it would be inappropriate to use them to obtain simple difference-in-mean earnings estimates of the effects of the Supported Work program. However, they may permit an unbiased regression estimation of program impacts within the context of a model that controls for key factors that affect earnings, but which neglects certain interactions and other second-order effects.

We drew two random comparison samples each for youth and for AFDC women: Stage 1 and Stage 2 random samples. The latter were obtained by subjecting CPS cases to a more extensive screening process for Supported Work eligibility than was used to obtain the Stage 1 random samples. Thus, the Stage 2 random samples should resemble Supported Work experimentals more closely than do the Stage 1 random samples.

1. Random Comparison Samples for Youth

The Stage 1 random comparison sample for youth was obtained by selecting 10 percent of the cases randomly from the 1976-1977 CPS general screen file, and then screening for age eligibility for the Supported Work

program. As indicated in Table III.1, 1,424 CPS cases remained at the end of this process.

The CPS Stage 2 comparison sample for youth was obtained by imposing an additional screen to the one used to obtain the Stage 1 sample. This screen excluded young people who were not high school dropouts or who were in school at the time they were interviewed. Although it would have been desirable to also limit this sample in other ways (for example, restricting it only to youth living in urban areas), the CPS sample of such cases is too small. Furthermore, limiting the Stage 2 Screens to a 10 percent random sample of the CPS youth general screen file would not have yielded a sufficient number of cases for the impact analysis. Thus, this more restrictive screen was applied to all cases in the CPS youth general screen file which satisfied the Stage 1 age criterion. As indicated in Table III.1, the resulting CPS Stage 2 youth random comparison sample consisted of 2,368 cases.

Distributions of CPS youth random comparison cases over values of variables believed to determine earnings are presented in Table III.2, as are those of Supported Work experimentals and controls. The Stage 1 and Stage 2 samples are distinctly different from the Supported Work samples with respect to gender, race, marital status, welfare receipt, and family income. However, the Stage 2 comparison sample resembles the Supported Work samples more closely with respect to education and occupation than does the Stage 1 sample.

The value of the CPS Stage 1 and Stage 2 samples as comparison groups for the Supported Work youth experimental sample can further be assessed by examining trends over time in average annual SSA earnings. An

TABLE III.1

DERIVATION OF RANDOM COMPARISON SAMPLES
FOR THE SUPPORTED WORK YOUTH SAMPLE

File/Screen	Number of Cases
A. CPS GENERAL SCREEN SAMPLE[a]	**116,245**
B. CPS YOUTH STAGE 1 SAMPLE (10 Percent Sample)	**1,424**
1. 10 Percent Random Sample of CPS General Screen Sample	11,624
2. Ages 16-20 in Interview Month	1,424
C. CPS YOUTH STAGE 2 SAMPLE (100 Percent Sample)	**2,368**
1. CPS General Screen Sample/Ages 16-20 in Interview Month	14,084
2. Ages 16-20/High School Dropout, Not Currently in School	2,368

NOTE: Screens were applied in the order in which they appear in this table.

[a]See Table II.8 for the derivation of the CPS general screen sample from the March 1976 and March 1977 CPS files.

ideal comparison group would be one whose average SSA earnings are similar to those of the Supported Work experimental and control groups during the pre-program years of 1972-1974.[1] Further, the average earnings of this ideal comparison group would be similar to those of the Supported Work control group during the in-program and post-program periods. As indicated in Figure III.1, the CPS Stage 1 and Stage 2 youth samples are inadequate Supported Work comparison groups according to both of these criteria (earnings data underlying Figure III.1 are presented in Appendix Table B.1). Average earnings for both comparison groups diverged sharply from those for Supported Work experimentals and controls during the pre-program period, and the gap between average comparison and control earnings widened over the years 1975-1979.

As indicated in Figure III.1, a comparison between the earnings of Supported Work youth experimentals and those of controls reveals a short-term positive net effect of the program. There is no evidence of a positive net effect when the earnings of Supported Work experimentals are compared with those of CPS Stage 1 and Stage 2 cases. Clearly, if these comparison groups are to be used to estimate the net impact of the Supported Work program on earnings, statistical techniques should be used to control for personal characteristics that determine the pattern of an individual's earnings over time.

---

[1] The experimental cases used in this study enrolled in the Supported Work program between April 1975 and July 1977.

TABLE III.2

PERCENT DISTRIBUTIONS OF CASES FROM THE SUPPORTED WORK YOUTH SAMPLES AND CPS YOUTH RANDOM COMPARISON SAMPLES

Characteristic	Supported Work		CPS Random Comparison Samples	
	Experimentals	Controls	Stage 1[a]	Stage 2[b]
Gender				
Male	88.7	86.9	48.8	48.4
Female	11.3	13.1	51.2	51.7
Age				
Younger than 19	61.0	56.6	58.9	60.4
19 or Older	37.1	41.9	41.1	39.6
Missing	1.9	1.5	0.0	0.0
Average	**17.8**	**18.0**	**18.1**	**18.1**
Race/Ethnicity				
Black, Non-Hispanic	72.4	69.9	10.8	12.0
Hispanic	17.8	19.6	6.4	9.3
White/Other	8.0	8.6	82.8	78.7
Missing	1.8	1.9	0.0	0.0
Education				
Fewer than 12 Years	99.3	99.4	56.5	100.0
12 or More Years	0.7	0.6	43.5	0.0
Missing	0.0	0.0	0.0	0.0
Average Years	**9.6**	**9.7**	**11.0**	**9.6**
Marital Status, Baseline Interview				
Married	4.6	3.5	12.6	28.5
Not Married	95.1	96.3	87.4	71.5
Missing	0.4	0.2	0.0	0.0
Occupation, Baseline Interview				
Sales/Clerical	10.6	11.2	17.4	12.5
Service	25.3	21.7	3.4	22.3
Professional	5.5	4.0	20.7	1.6
Other	27.2	29.5	28.8	33.9
Missing/None	31.4	33.6	29.7	29.7
Public Sector Employment, Baseline Interview				
Yes	n.a.	n.a.	6.6	4.9
No			93.4	95.1

TABLE III.2 (continued)

Characteristic	Supported Work		CPS Random Comparison Samples	
	Experimentals	Controls	Stage 1[a]	Stage 2[b]
Percent of Time Employed, Base Year				
0	41.5	43.7	29.6	29.7
1 - 49	44.5	40.4	31.0	29.9
50 - 89	9.9	12.1	17.1	18.9
≥ 90	1.4	1.8	22.3	21.5
Missing	2.7	2.1	0.0	0.0
SSA Earnings, Base Year[c]				
$0	1.4	0.0	32.9	30.7
$1 - $1,000	84.5	87.6	25.5	26.8
$1,001 - $5,000	14.1	12.4	33.4	34.5
> $5,000	0.0	0.0	8.3	8.1
Average	**$578**	**$550**	**$1500**	**$1528**
Change in SSA Earnings, Pre-Base Year to Base Year[c]				
Negative	30.0	34.2	17.8	19.7
Zero[d]	10.3	12.2	29.6	25.8
Positive	59.7	53.5	52.6	54.6
Average Change	**$126**	**$98**	**$563**	**$589**
Welfare Receipt, Base Year				
Yes	24.6	24.6	2.0	6.5
No	74.2	73.8	98.0	93.5
Missing	1.2	1.6	0.0	0.0
Average Monthly Family Income, Base Year	**$92**	**$95**	**$346**	**$329**
Number of Cases	566	678	1424	2368

[a]This sample includes a randomly selected 10 percent of the 16-to-20 year olds in the CPS sample.

[b]This sample includes all 16-to-20 year old school dropouts in the CPS sample.

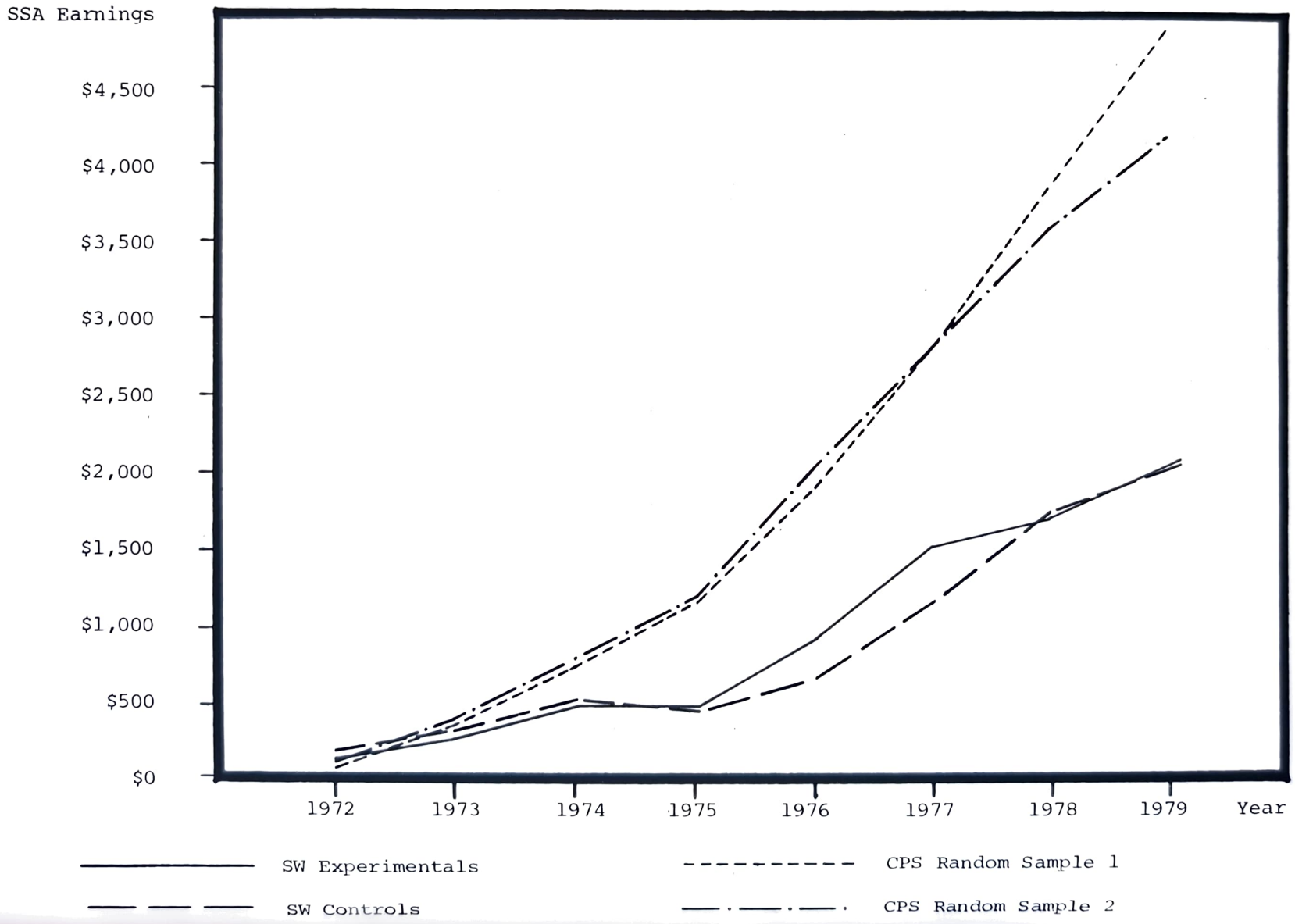
[c]SSA earnings for Supported Work cases are group mean values rather than values for individuals. SSA earnings for CPS cases are values for individuals.

[d]The zero change in the SSA earnings category includes cases whose SSA earnings did not change at all, 5 percent of the smallest negative-change cases, and 5 percent of the smallest positive-change cases.

n.a. means not available.

FIGURE III.1

AVERAGE SSA EARNINGS FOR YOUTH: SUPPORTED WORK EXPERIMENTALS AND CONTROLS AND CPS RANDOM SAMPLES 1 AND 2



2. Random Comparison Samples for AFDC Recipients

Stage 1 and Stage 2 random comparison samples were also selected for AFDC women. Because the Stage 1 screens were more restrictive for AFDC women than for youth, all of the cases from the 1976-1977 CPS AFDC general screen file were used. As indicated in Table III.3, the successive application of screens with respect to age, sex, age of youngest dependent, and welfare receipt reduced the CPS AFDC general screen file to 1,995 Stage 1 cases. Women with dependents younger than age 6 were excluded from the AFDC phase of the Supported Work program. This additional restriction was imposed on the Stage 1 CPS AFDC comparison sample, leaving 909 Stage 2 comparison cases.[1]

As indicated in Table III.4, the Supported Work samples and the CPS comparison samples of AFDC women do not correspond very closely on a number of characteristics. The comparison samples contain relatively more white women and women who were employed at least 50 percent of the available time during the base year. Women in these samples also had higher base-year family incomes. The screen on dependents younger than age 6 created a Stage 2 comparison sample that more closely resembles the Supported Work samples with respect to the age of the youngest child than does the Stage 1 comparison sample. However, this screen pushed the average base-year age of Stage 2 women up to 37 years, as compared with 32 years for Stage 1 women and 33 for the Supported Work sample.

---

[1] Although, it would have been desirable to impose a screen on long-term welfare dependence, data to do this are not available on the CPS. Restricting the sample in other ways (for example, to urban residents) would have left too few observations to conduct the analysis.

TABLE III.3

DERIVATION OF RANDOM COMPARISON SAMPLES
FOR THE SUPPORTED WORK AFDC SAMPLE

File/Screen	Number of Cases
A. CPS GENERAL SCREEN SAMPLE[a]	**116,245**
B. CPS AFDC STAGE 1 SAMPLE (100 Percent Sample)	**1,995**
1. CPS General Screen Sample/Ages 18-64 in Interview Month	111,566
2. Ages 18-64/Female	59,674
3. Ages 18-64/Female/Dependents Younger than age 18	32,435
4. Ages 18-64/Female/Dependents Younger than Age 18/Welfare Recipient Base Year	1,995
C. CPS AFDC STAGE 2 SAMPLE (100 Percent Sample)	**909**
1. CPS AFDC Stage 1 Sample	1,995
2. No Dependents Younger than Age 6	909

NOTE: Screens were applied in the order in which they appear in this table.

[a]See Table II.8 for the derivation of the CPS general screen sample from the March 1976 and March 1977 CPS files.

The average earnings of CPS AFDC Stage 1 and Stage 2 comparison cases were generally higher than those of Supported Work AFDC cases during the pre-program period (Figure III.2); however, there is no evidence of dramatically different growth rates. More importantly, the average earnings of the Stage 1 and Stage 2 groups do not diverge systematically from those of the Supported Work AFDC control group during the in-program and post-program periods. By 1978, the earnings advantage of comparison-group women relative to control-group women had stabilized at several hundred dollars. We suspect that, as indicated by earnings patterns, the relative success of the AFDC random comparison groups is attributable largely to the Stage 1 screen for base-year receipt of welfare benefits.

A comparison between the earnings of Supported Work AFDC experimentals and those of controls reveals a large positive effect of the program in 1977. This effect is partially attributable to in-program earnings. Program participants continued to earn more than their control-group counterparts in 1978 and 1979, but the differences were much smaller due to the reduced incidence of in-program earnings.

A comparison between the earnings of Supported Work AFDC experimentals and those of the CPS Stage 1 and Stage 2 random comparison cases reveals a much smaller positive short-run effect of the Supported Work program and a negligible longer-run effect. These simple net-impact estimates do not include adjustments for differences between the experimental and comparison cases with respect to factors that affect earnings. Matching procedures (described in the remainder of this chapter) and econometric modeling (discussed in Chapter IV) were used to obtain better estimates of net program impacts from experimental and comparison samples.

TABLE III.4

PERCENT DISTRIBUTION OF CASES FROM THE SUPPORTED WORK AFDC SAMPLES AND CPS AFDC RANDOM COMPARISON SAMPLES

Characteristic	Supported Work		CPS Random Comparison Samples	
	Experimentals	Controls	Stage 1[a]	Stage 2[b]
Age				
Younger than 21	1.0	0.5	9.4	2.9
21 - 34	58.1	59.0	56.5	37.5
35 or Older	40.8	40.0	34.1	59.6
Missing	0.1	0.5	0.0	0.0
Average	**33.3**	**33.5**	**31.6**	**37.5**
Race/Ethnicity				
Black, Non-Hispanic	83.5	81.6	37.8	37.0
Hispanic	12.0	13.0	12.0	11.6
White/Other	4.5	5.5	50.2	51.5
Education				
Fewer than 12 Years	69.7	69.2	59.5	62.6
12 or More Years	30.3	30.8	40.5	37.4
Missing	0.0	0.0	0.0	0.0
Average Years	**10.3**	**10.3**	**10.1**	**9.7**
Occupation, Baseline Interview				
Sales/Clerical	7.3	8.0	11.0	11.7
Service	10.3	8.4	16.6	18.4
Professional	1.2	1.6	2.2	2.2
Other	6.7	7.7	8.6	8.9
Missing/None	74.5	74.3	61.6	58.9
Public Sector Employment, Baseline Interview				
Yes	n.a.	n.a.	6.2	6.5
No			93.8	93.5
Percent of Time Employed, Base Year				
0	80.2	78.2	61.6	58.9
1 - 49	15.1	16.8	16.7	14.0
50 - 89	2.9	3.4	10.9	12.8
$\geq$ 90	1.4	0.7	10.8	14.4
Missing	0.4	0.9	0.0	0.0

TABLE III.4 (continued)

Characteristic	Supported Work		CPS Random Comparison Samples	
	Experimentals	Controls	Stage 1[a]	Stage 2[b]
SSA Earnings, Base Year[c]				
$0	10.6	5.4	59.6	58.5
$1 - $1,000	79.4	84.0	15.6	12.7
$1,001 - $5,000	10.0	10.6	19.9	23.0
≥ $5,000	0.0	0.0	4.9	5.8
Average	**$458**	**$463**	**$919**	**$1100**
Change in SSA Earnings, Pre-Base Year to Base Year[c]				
Negative	47.9	56.2	25.1	23.9
Zero[d]	16.2	12.3	49.6	50.4
Positive	35.9	31.4	25.3	25.7
Average Change	**$-55**	**$-153**	**$39**	**$32**
Welfare Receipt, Base Year				
Yes	99.9	99.2	100.0	100.0
No	0.0	0.3	0.0	0.0
Missing	0.1	0.5	0.0	0.0
Average Monthly Family Income, Base Year	**$310**	**$309**	**$554**	**$594**
Age of Youngest Child, Baseline Interview				
Younger than age 6	2.1	2.6	54.4	0.0
6 - 12 Years	71.1	72.3	30.2	66.3
12 or Older	25.6	23.9	15.3	33.7
Missing	1.2	1.8	0.0	0.0
Number of Cases	800	802	1995	909

[a]This sample includes women age 16 to 24 with one or more children who receive welfare in the year preceding the CPS interview.

[b]This sample includes all women in the State 1 sample (see footnote a) who have no child younger than six.

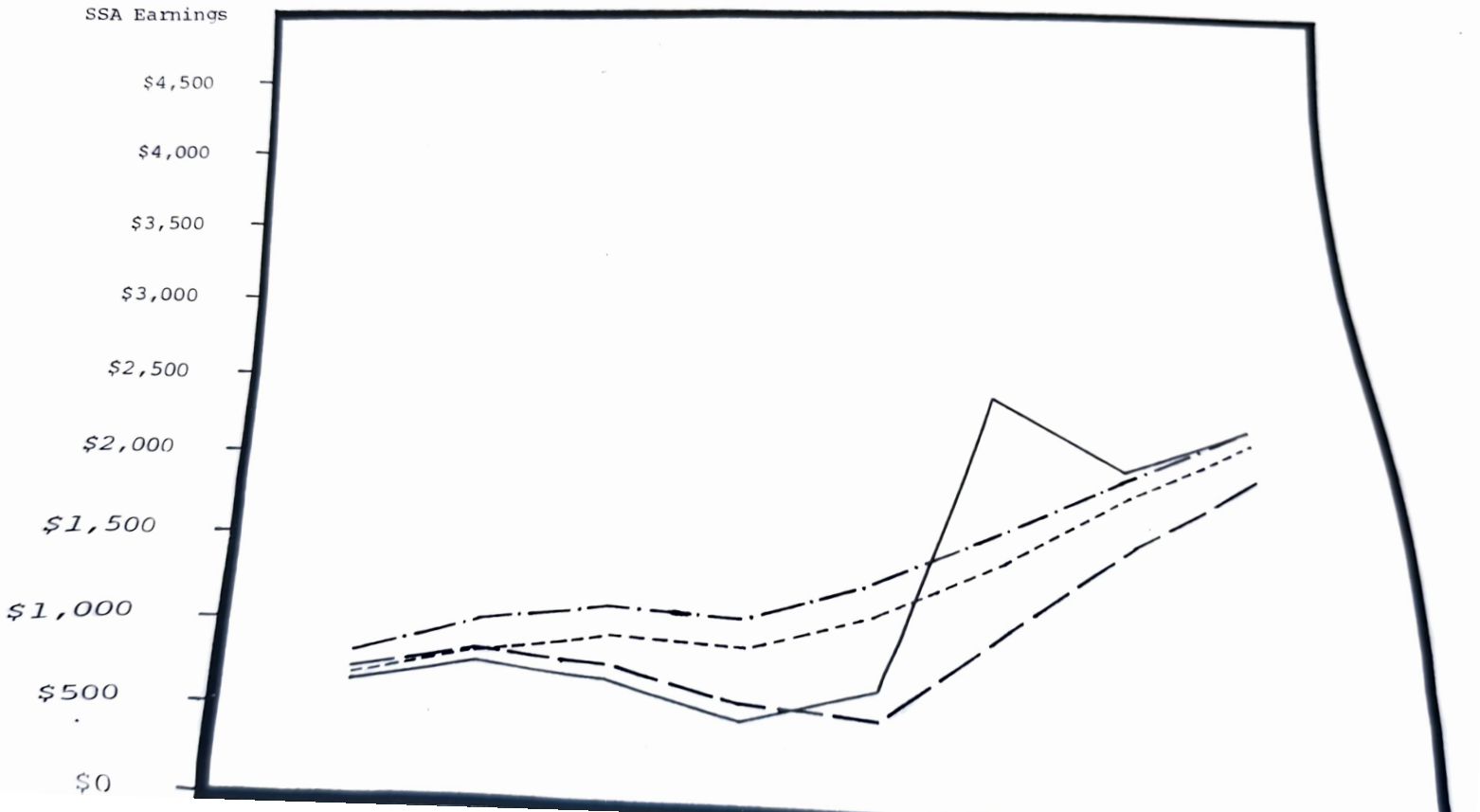
[c]SSA earnings for Supported Work cases are group mean values rather than values for individuals. SSA earnings for CPS cases are values for individuals.

[d]The zero change in SSA earnings category includes cases whose SSA earnings did not change at all, 5 percent of the smallest negative-change cases, and 5 percent of the smallest positive-change cases.

n.a. means not available.

FIGURE III.2

AVERAGE SSA EARNINGS FOR AFDC WOMEN: SUPPORTED WORK EXPERIMENTALS AND CONTROLS AND CPS RANDOM SAMPLES 1 AND 2



C. STRATIFIED MATCHING

The objective of this project was to investigate procedures for obtaining comparison groups that could be used to obtain reliable estimates of program effects without extensively modeling earnings. As we have shown, the random selection of CPS comparison cases after minimal screening is deficient in this respect; the simple difference between the mean earnings of Supported Work experimentals and CPS random comparison cases appears to be an unreliable estimator of the net impact of the Supported Work program. In this section, we describe a stratified matching procedure that was used to obtain CPS comparison samples for the Supported Work youth and AFDC experimental samples. A preliminary assessment of the utility of the matched CPS samples as comparison groups was made by examining their characteristics and earnings patterns relative to those of Supported Work experimentals and controls.

Stratified matching (cell matching) requires identifying a group of match variables that are significant predictors of the outcome of interest--in this case earnings. Comparison cases that resemble experimental cases with respect to these variables are then selected. We used stepwise regression analysis to obtain two different groups of match variables and to rank them by the proportion of the variation in earnings they explain.

1. Stratified Matching for Youth

We wished to investigate whether comparison samples based on detailed matching procedures yielded better net-impact estimates than comparison samples based on simple matching procedures. Consequently, short and long lists of potential match variables were subjected to a stepwise regression analysis of their power to explain earnings. Base-year

measures of education, race, and sex constituted the short list. The long list included additional variables that measured age, family income, past earnings, and base-year employment. These two groups of variables were used to analyze base-year interview earnings and 1979 SSA earnings separately for the CPS youth Stage 2 sample and the Supported Work experimental sample.[1] The order in which these variables entered the stepwise regressions is shown in Table III.5. Because we were ultimately interested in estimating the longer-run impacts of the Supported Work program, we based our stratified matching on rankings produced by the analysis of the 1979 SSA earnings of CPS cases.

Based on these rankings, both cells and priorities for cell collapsing were defined. In all cases, we required the presence of four or more Supported Work and CPS cases in each cell. As a result, the last three of the nine long-list match variables to enter the stepwise earnings regression were not used in the stratified matching procedure, since the samples were too small to enable us to match at this level of detail and to meet our minimum cell requirement.

Categories of the six active long-list match variables are shown in Table III.6. A six-digit number uniquely identifies each of the 320 cells defined by categories of these variables. From left to right, the digits represent categories of match variables in the order of their selection into the stepwise earnings equation. The majority of the cells had fewer than four Supported Work and CPS cases. Such cells were collapsed (categories were combined) until more broadly defined cells that contained at

---

[1] The experimental sample rather than the control group sample was used since in cases where comparison group methods are needed and used, control groups are unavailable.

TABLE III.5

ORDER IN WHICH POTENTIAL MATCH VARIABLES ENTERED STEPWISE EARNINGS REGRESSIONS FOR YOUTH

	Short Match List				Long Match List			
	CPS[a]		Supported Work		CPS[a]		Supported Work	
Variable	Base Year Earnings[b]	1979 SSA $	Base Year Earnings[b]	1979 SSA $	Base Year Earnings[b]	1979 SSA $	Base Year Earnings[b]	1979 SSA $
Black, Non-Hispanic	2	2	3	1	-	4	-	2
Hispanic	-	-	-	-	-	-	-	-
Male	1	1	2	-	5	1	5	-
≤ 8 Years of School	3	3	1	-	-	5	-	-
Ages 16-18	n.a.	n.a.	n.a.	n.a.	6	6	4	-
Middle Per Capita Family Income	n.a.	n.a.	n.a.	n.a.	-	-	-	-
High Per Capita Family Income	n.a.	n.a.	n.a.	n.a.	4	9	3	-
> 0 But < 26 Weeks of Employment in Base Year	n.a.	n.a.	n.a.	n.a.	7	8	2	-
≥ 26 Weeks of Employment in Base Year	n.a.	n.a.	n.a.	n.a.	3	7	1	-
SSA Earnings, Pre-Base Year	n.a.	n.a.	n.a.	n.a.	1	2	-	1
Change in SSA Earnings, Pre-Base Year to Base Year	n.a.	n.a.	n.a.	n.a.	2	3	-	3

NOTE: A dash (-) indicates that a variable did not enter an equation with a significance level of 0.15 or less.

[a]The CPS sample used was the youth Stage 2 sample described in Table III.1.

[b]The base-year earnings variable was interview earnings, rather than SSA earnings.

n.a. means variable was not considered for inclusion in the equation.

TABLE III.6

CATEGORIES OF VARIABLES USED IN STRATIFIED MATCHING FOR YOUTH

Variable	Categories
Male*	0 - No 1 - Yes
SSA Earnings, Pre-Base Year	1 - Low 2 - Medium 3 - High[a,b] 4 - Very High[a,b,c]
Change in SSA Earnings, Base Year -1 to Base Year	0 - Very Negative[b] 1 - Negative[b] 2 - Zero 3 - Positive[b] 4 - Very Positive[b]
Black, Non-Hispanic*	0 - No 1 - Yes
$\leq$ 8 Years of School*	0 - No 1 - Yes
Age 16 - 18	0 - No 1 - Yes

[a] These categories were combined in Variant B of the long-list cell match.

[b] These categories were combined in Variant C of the long-list cell match.

[c] CPS cases only.

*Variable appeared in both the short and long match lists.

least four cases of each type were obtained. When collapsing was necessary, the appropriate digits in the number representing the new cell were filled with values of "9."

Combining categories reduced the number of cells defined by the long list of match variables to 34. The distributions of Supported Work experimental and CPS youth Stage 2 cases across these cells are shown in Table III.7. Note that three variants of the matched CPS youth sample were obtained. To obtain Variant A, we used the full range of categories for SSA earnings and changes in SSA earnings. For example, CPS cases with "very high" SSA earnings one year prior to the base year were not matched with Supported Work cases with "high" earnings, and they were consequently excluded from CPS match Variant A. In obtaining Variant B, we did not use the distinction between "high" and "very high" SSA earnings one year prior to the base year, which led to the selection of many more CPS cases but which reduced the precision of the match. With Variant C, we went one step beyond Variant B, eliminating the "very negative"/"negative" and "very positive"/"positive" distinctions for the change in SSA earnings between the first year prior to the base year and the base year. This again increased the size of the matched comparison sample, but further decreased the quality of the match.

In selecting matched comparison groups A, B, and C, it was unnecessary to collapse categories for the first three long-list variables to enter the stepwise earnings regression. However, it was necessary to ignore many distinctions among categories of lower-order variables in order to obtain CPS matches for Supported Work cases. Table III.8 shows the extent to which it was necessary to collapse categories for each of the six

TABLE III.7

SUMMARY OF LONG LIST CELL MATCHING RESULTS
FOR THE SUPPORTED WORK YOUTH EXPERIMENTAL SAMPLE

Cell Number	Supported Work Youth Sample		CPS Youth Long List Matched Samples					
			Variant A[a]		Variant B[b]		Variant C[c]	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
1. 012199	8	1.5	72	8.3	12	1.1	72	6.0
2. 013099	4	0.7	129	14.8	129	11.5	129	10.8
3. 013199	13	2.4	16	1.8	28	2.5	28	2.3
4. 021999	5	0.9	62	7.1	62	5.5	62	5.2
5. 023099	5	0.9	29	3.3	70	6.3	70	5.8
6. 023199	10	1.8	4	0.5	4	0.4	4	0.3
7. 031999	8	1.5	28	3.2	76	6.8	76	6.3
8. 032999	5	0.9	5	0.6	5	0.4	5	0.4
9. 033999	6	1.1	28	3.2	89	7.9	89	7.4
10. 111999	8	1.5	14	1.6	14	1.3	14	1.2
11. 112999	28	5.1	191	21.9	191	17.1	191	15.9
12. 113009	14	2.5	86	9.9	86	7.7	86	7.2
13. 113010	4	0.7	4	4.6	12	1.1	12	1.0
14. 113011	17	3.1	10	1.1	24	2.1	24	2.0
15. 113100	13	2.4	11	1.3	11	1.0	11	0.9
16. 113101	39	7.1	11	1.3	11	1.0	11	0.9
17. 113119	34	6.2	4	0.5	4	0.4	4	0.3
18. 121099	7	1.3	26	3.0	26	2.3	26	2.2
19. 121199	40	7.3	4	0.5	4	0.4	4	0.3
20. 122999	11	2.0	7	0.8	7	0.6	7	0.6
21. 123009	10	1.8	25	2.9	62	5.6	62	5.2
22. 123019	20	3.6	6	0.7	6	0.5	6	0.5
23. 123109	49	8.9	6	0.7	6	0.5	6	0.5
24. 123119	27	4.9	6	0.7	6	0.5	6	0.5
25. 131009	11	2.0	12	1.4	34	3.0	34	2.8
26. 131019	16	2.9	5	0.1	8	0.7	21	1.8
27. 131100	23	4.2	9	1.0	9	0.8	9	0.8
28. 131101	22	4.0	4	0.5	5	0.4	9	0.8
29. 131119	19	3.5	5	0.6	5	0.4	5	0.4
30. 132999	17	3.1	5	0.6	5	0.4	12	1.0
31. 133099	18	3.3	28	3.2	28	2.5	73	6.1
32. 133100	16	2.9	4	0.5	6	0.5	17	1.4
33. 133101	12	2.2	8	0.9	8	0.7	8	0.7
34. 133119	11	2.0	7	0.8	7	0.6	7	0.6
Total	550	100.0	871	100.0	1,120	100.0	1,200	100.0

NOTE: Initially, all Supported Work and CPS sample members were assigned to cells defined by unique combinations of the six characteristics shown in Table III.6. Cells were then collapsed so as to achieve cells with a minimum of four cases from each sample. Collapsing occurred across characteristics in the inverse order of their presentation in Table III.6. A value of "9" in a cell number indicates that cells were collapsed on the attribute denoted by that digit.

[a]This variant included the full range of SSA earnings categories listed in Table III.6.

[b]This variant collapsed the categories of high and very high positive SSA earnings in the year prior to the base year for CPS observations.

[c]This variant collapsed the categories of very negative and negative and the categories of very positive and positive changes in SSA earnings, in addition to collapsing the high and very high categories of SSA earnings for the CPS observations.

TABLE III.8

NUMBER OF CASES IN COLLAPSED CELLS FROM THE SUPPORTED WORK YOUTH EXPERIMENTAL SAMPLE AND THREE MATCHED CPS SAMPLES, USING THE LONG LIST OF MATCH VARIABLES

	Supported Work Youth Sample		CPS Long List Matched Samples					
			Variant A		Variant B		Variant C	
Variable	Exact Match	In Collapsed Cells	Exact Match	In Collapsed Cells	Exact Match	In Collapsed Cells	Exact Match	In Collapsed Cells
Gender	550	0	871	0	1,120	0	1,200	0
SSA Earnings, Base Year -1	550	0	871	0	1,120	0	1,200	0
Change in SSA Earnings, Base Year -1 to Base Year	550	0	871	0	1,120	0	1,200	0
Black, Non-Hispanic	462	88	531	340	671	449	744	456
$\leq$ 8 Years of School	357	193	225	646	370	750	338	862
Age 16-18	146	404	63	808	146	974	101	1,099

NOTE: The CPS sample used was the Youth Stage 2 sample described in Table III.1.

active long-list match variables so as to achieve successful matches of Supported Work experimentals and CPS Stage 2 cases for youth.

Stratified matching of CPS youth Stage 2 cases with Supported Work youth experimental cases based on the short list of match variables (gender, race, and education) did not necessitate that cells be collapsed. None of the original eight cells defined by categories of these variables included less than four Supported Work or CPS cases. The distributions of Supported Work and CPS cases across these cells are shown in Table III.9.

Characteristics of the CPS short- and long-list matched samples for youth are shown in Table III.10. In all cases, the sample characteristics have been weighted by the ratio of Supported Work to CPS cases in the stratified match cell. This maximizes the size of the comparison sample at the same time it preserves the similarity of the groups on key characteristics.

Variants A-C of the CPS long-list match exhibit very similar characteristics. These comparison samples resemble the Supported Work youth samples much more closely than do the Stage 1 and Stage 2 random comparison samples--particularly with respect to such important match variables as gender, race, and change in SSA earnings. The short-list matched sample corresponds slightly more with the Supported Work samples on gender and race than do the long-list matched samples, since the short-list matching process did not necessitate that categories of these two match variables be collapsed. This sample does not correspond with the Supported Work samples as closely as the long-list variants A-C for variables not included in the short match list. However, even for marital status and

TABLE III.9

SUMMARY OF SHORT LIST CELL MATCHING RESULTS FOR THE SUPPORTED WORK YOUTH EXPERIMENTAL SAMPLE

Cell Number[a]	Supported Work Youth Sample		CPS Youth Short List Matched Sample	
	Number	Percent	Number	Percent
1. 000	8	1.4	917	38.8
2. 001	7	1.3	155	6.6
3. 010	41	7.4	136	5.8
4. 011	8	1.4	13	0.6
5. 100	47	8.5	849	35.9
6. 101	84	15.1	160	6.8
7. 110	241	43.3	102	4.3
8. 111	120	21.6	33	1.4
Total	556	100.0	2,365	100.0

[a]The first digit of the cell number indicates categories of the variable "Male," the second digit indicates categories of the variable "Black, Non-Hispanic," and the third digit indicates categories of the variable "$\leq$ 8 Years of School." Codes for the categories of these variables are provided in Table III.6.

TABLE III.10

PERCENT DISTRIBUTIONS OF CASES FROM THE SUPPORTED WORK YOUTH SAMPLES AND WEIGHTED DISTRIBUTIONS OF MATCHED CPS YOUTH CASES FROM THE SHORT LIST MATCH AND LONG LIST MATCHES A, B, AND C

Characteristic	Supported Work		CPS Comparison Groups			
	Experimentals	Controls	Short List	Match A	Match B	Match C
Gender						
Male	88.7	86.9	88.5	88.4	88.4	88.4
Female	11.3	13.1	11.5	11.6	11.6	11.6
Age						
Younger than 19	61.0	56.6	54.6	61.2	61.2	59.2
19 or Older	37.1	41.9	45.4	38.8	38.8	40.8
Missing	1.9	1.5	0.0	0.0	0.0	0.0
Average	**17.8**	**18.0**	**18.2**	**18.0**	**18.0**	**18.1**
Race/Ethnicity						
Black, Non-Hispanic	72.4	69.9	73.7	58.6	58.6	58.4
Hispanic	17.8	19.6	3.5	3.7	4.5	4.4
White/Other	8.0	8.6	22.8	37.7	36.9	37.2
Missing	1.8	1.9	0.0	0.0	0.0	0.0
Education						
Fewer than 12 Years	99.3	99.4	100.0	100.0	100.0	100.0
12 or More Years	0.7	0.6	0.0	0.0	0.0	0.0
Missing	0.0	0.0	0.0	0.0	0.0	0.0
Average Years	**9.6**	**9.7**	**8.6**	**9.2**	**9.2**	**9.2**
Marital Status, Baseline Interview						
Married	4.6	3.5	13.9	9.5	9.7	11.7
Not Married	95.1	96.3	86.1	90.5	90.3	88.3
Missing	0.4	0.2	0.0	0.0	0.0	0.0
Occupation, Baseline Interview						
Sales/Clerical	10.6	11.2	3.5	6.9	6.7	6.3
Service	25.3	21.7	16.7	23.7	23.6	22.5
Professional	5.5	4.0	0.5	0.2	0.3	0.3
Other	27.2	29.5	39.5	41.1	41.5	42.8
Missing/None	31.4	33.6	54.8	28.1	27.9	28.1
Public Sector Employment, Baseline Interview						
Yes	n.a.	n.a.	10.1	10.1	10.2	10.4
No			89.9	89.9	89.8	89.6

TABLE III.10 (continued)

Characteristic	Supported Work		CPS Comparison Groups			
	Experimentals	Controls	Short List	Match A	Match B	Match C
Percent of Time Employed, Base Year						
0	41.5	43.7	87.3	89.9	89.3	87.2
1 - 49	44.5	40.4	12.7	10.1	10.7	12.8
50 - 89	9.9	12.1	0.0	0.0	0.0	0.0
≥ 90	1.4	1.8	0.0	0.0	0.0	0.0
Missing	2.7	2.1	0.0	0.0	0.0	0.0
SSA Earnings, Base Year[a]						
$0	1.4	0.0	31.8	15.6	16.9	16.5
$1 - $1,000	84.5	87.6	26.0	47.8	42.3	40.1
$1,001 - $5,000	14.1	12.4	34.3	35.2	39.4	39.9
≥ $5,000	0.0	0.0	7.9	1.3	1.4	3.5
Average	**$578**	**$550**	**$1508**	**$945**	**$1027**	**$1267**
Change in SSA Earnings, Pre-Base Year to Base Year[a]						
Negative	30.0	34.2	21.6	28.9	28.9	28.9
Zero[b]	10.3	12.2	29.4	10.4	10.4	11.1
Positive	59.7	53.5	49.0	60.7	60.7	60.0
Average Change	**$126**	**$98**	**$484**	**$360**	**$442**	**$432**
Welfare Receipt--Base Year						
Yes	24.6	24.6	5.2	5.3	5.3	5.4
No	74.2	73.8	94.8	94.7	94.7	94.6
Missing	1.2	1.6	0.0	0.0	0.0	0.0
Average Monthly Family Income, Base Year	**$92**	**$95**	**$402**	**$324**	**$338**	**$360**
Number of Cases	566	678	2365	871	1120	1120

[a]SSA earnings for Supported Work cases are group mean values rather than values for individuals. SSA earnings for CPS cases are values for individuals.

[b]The zero change in the SSA earnings category includes cases whose SSA earnings did not change at all, 5 percent of the smallest negative-change cases, and 5 percent of the smallest positive-change cases.

n.a. means not available.

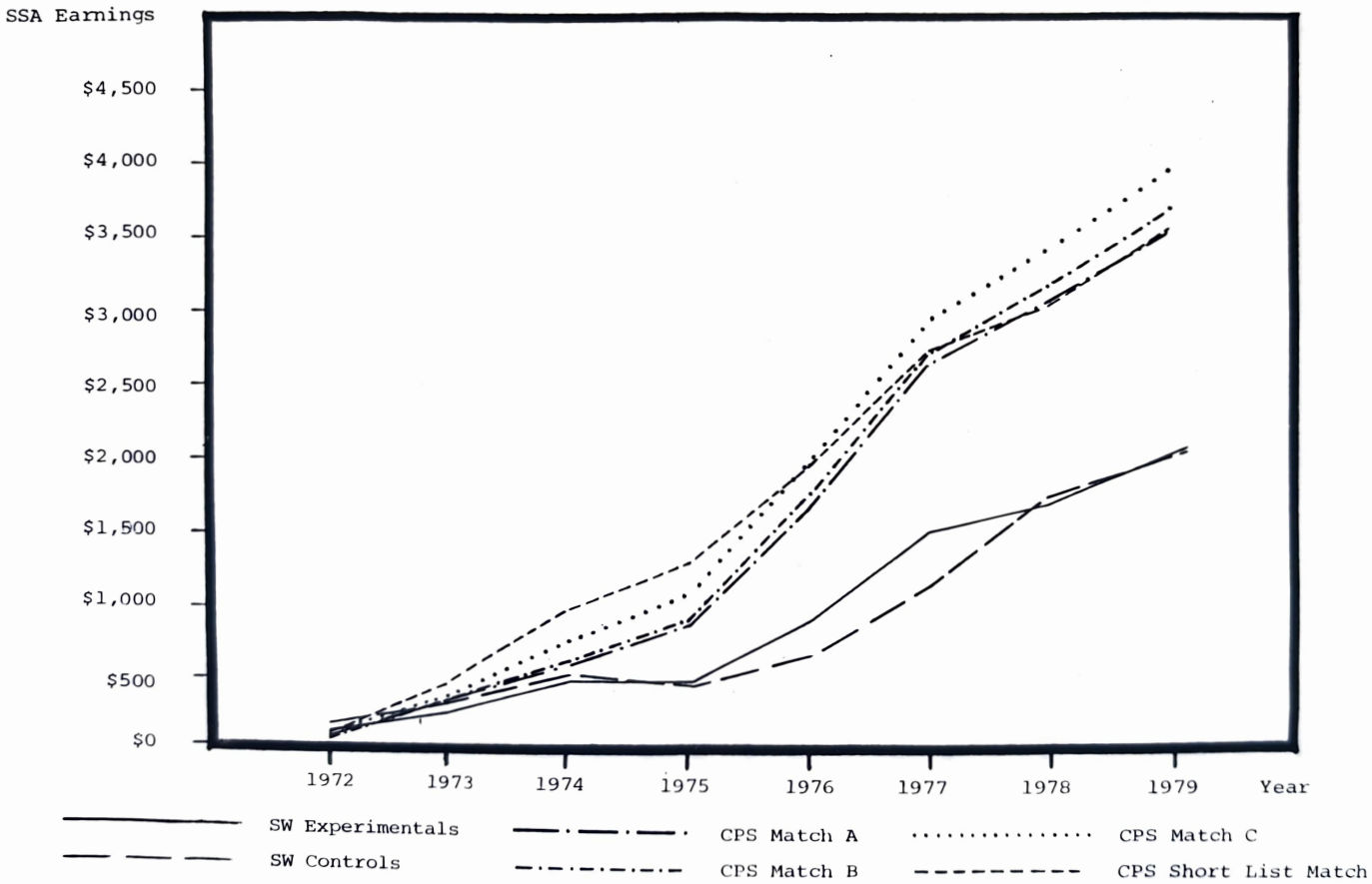
other variables that were in neither match list, the matched samples resemble the Supported Work samples more closely than do the Stage 1 and Stage 2 random comparison samples.

Of the four matched samples for youth, the one based on the short list of match variables had mean SSA earnings for the pre-program period that differed most from those of the Supported Work samples. This was expected, since measures of pre-program earnings were included in the long list of match variables but not in the short list. As indicated by Figure III.3, the mean earnings of the short-list match group in the longer-run were quite similar to those of the long-list match groups. Earnings for all four groups diverged sharply from those of Supported Work controls during the in-program and post-program periods; however, the divergence was smaller than for the CPS youth random comparison samples (Figure III.1).

Of the long-list match groups, Variant A was based on the finest categorization of pre-program earnings. The effect of doing so is apparent in Figure III.3, which shows that the pre-program earnings of Variant A cases equaled those of Supported Work cases more closely than did the earnings of Variant B and C cases. This correspondence was much closer than for the random comparison cases. However, the longer-run earnings of Variant A cases differed fundamentally from those of Supported Work controls, despite the high degree of correspondence between pre-program earnings and the matching of Variant A cases with Supported Work experimentals on the basis of personal characteristics known to affect earnings in the longer run. A simple comparison of the mean earnings of Variant A cases and Supported Work experimentals reveals large short-run and longer-run negative effects of the Supported Work program on youth. We have

FIGURE III.3

AVERAGE SSA EARNINGS FOR YOUTH: SUPPORTED WORK EXPERIMENTALS AND CONTROLS; WEIGHTED CPS MATCHES A, B, AND C; AND WEIGHTED CPS SHORT LIST MATCH



previously noted that a comparison of the mean earnings of Supported Work experimentals and controls reveals a short-run positive effect of the program and a longer-run zero effect.

It is not surprising that eligible youth who applied to the Supported Work program exhibited very different earnings patterns over time than did eligible youth who did not apply but who were similar with respect to age, race, sex, and education. The former are likely to have differed systematically from the latter with respect to unobservable factors that determine earnings. We had hoped that by matching on pre-program earnings we would have reduced such unobservable differences between matched Supported Work and CPS samples. However, while it may have reduced unobservable differences somewhat, Figure III.3 shows that matching on pre-program earnings did not create youth comparison groups that resembled the true control group with respect to long-run earnings. We speculate that the young ages of the target individuals may have contributed to our inability to obtain matched comparison cases with long-run earnings similar to those of the control cases. Recent past earnings may be less reliable predictors of future earnings for persons in their late teens and early twenties than for older people, who typically have more work experience and more stable earnings patterns.

2. Stratified Matching for AFDC Recipients

We also used stratified matching to select cases from the CPS Stage 2 AFDC file that resembled Supported Work AFDC experimentals. As with youth, we used a stepwise regression analysis of earnings to identify match variables and the order of cell collapsing. Variables available for selection into the earnings regressions are shown in Table III.11. They

TABLE III.11

ORDER IN WHICH POTENTIAL MATCH VARIABLES ENTERED STEPWISE EARNINGS REGRESSIONS FOR AFDC WOMEN

Variable	Short Match List				Long Match List			
	CPS[a]		Supported Work		CPS[a]		Supported Work	
	Base Year Earnings[b]	1979 SSA $	Base Year Earnings[b]	1979 SSA $	Base Year Earnings[b]	1979 SSA $	Base Year Earnings[b]	1979 SSA $
Black, Non-Hispanic	2	2	-	1	-	5	-	4
Hispanic	-	-	-	-	-	-	-	-
≥ 12 Years of School	1	1	-	-	5	-	-	-
No Dependents < 12 Years Old	-	-	-	-	-	-	-	-
Age < 35	n.a.	n.a.	n.a.	n.a.	-	4	-	5
Middle Per Capita Family Income	n.a.	n.a.	n.a.	n.a.	-	-	3	-
High Per Capita Family Income	n.a.	n.a.	n.a.	n.a.	4	-	2	1
> 0 Weeks of Employment in Base Year	n.a.	n.a.	n.a.	n.a.	1	3	1	-
SSA Earnings, Pre-Base Year	n.a.	n.a.	n.a.	n.a.	2	1	-	2
Change in SSA Earnings, Pre-Base Year to Base Year	n.a.	n.a.	n.a.	n.a.	3	2	-	3

NOTE: A dash (-) indicates that a variable did not enter an equation with a significance level of 0.15 or less.

[a]The CPS sample used was the AFDC Stage 2 sample described in Table III.3.

[b]The base-year earnings variable was interview earnings, rather than SSA earnings.

n.a. means variable was not considered for inclusion.

differ from those available for selection into the stepwise regressions for youth (Table III.5) primarily with respect to measures of sex, age, education, and the presence of dependents. The short- and long-list cell matches were based on stepwise regression results for CPS cases, using 1979 SSA earnings as the dependent variable. Corresponding regression results for Supported Work cases were similar to these.

Categories of variables which comprised the long and short AFDC match lists are shown in Table III.12. These categories define 160 long-list cells and 4 short-list cells. During the matching process, these categories were consolidated into, respectively, 25 cells and 3 cells. Tables III.13 and III.14 provide information on the long-list match of Supported Work experimentals and CPS Stage 2 cases. In obtaining CPS matches for the Supported Work AFDC experimentals, it was unnecessary to collapse across any categories of the two most important long-list match variables. Minimal collapsing on "Age < 35" and extensive collapsing on variables that measured base-year employment and race were required. Given the relatively large proportion of CPS Stage 2 AFDC cases that were matched (554 of 909, or 61 percent),[1] we did not obtain alternative long-list matches based on collapsed extreme categories of pre-program SSA earnings.

Although the short match list includes only two variables (dichotomous measures of education and race), some cell collapsing was nonetheless necessary to achieve an acceptable match. Racial distinctions were ignored in matching Supported Work cases that had twelve or more years of education with CPS counterparts. Distributions of matched Supported

---

[1] Only 37 percent of the CPS Stage 2 youth cases were included in Variant A of the long-list cell match for youth.

TABLE III.12

CATEGORIES OF VARIABLES USED IN STRATIFIED MATCHING FOR AFDC WOMEN

Variable	Categories
SSA Earnings, Pre-Base Year	1 - Low 2 - Medium 3 - High 4 - Very High[a]
Change in SSA Earnings Pre-Base Year to Base Year	0 - Very Negative 1 - Negative 2 - Zero 3 - Positive 4 - Very Positive
Age < 35	0 - No 1 - Yes
> 0 Weeks of Employment in Base Year	0 - No 1 - Yes
Black, Non-Hispanic*	0 - No 1 - Yes
≥ 12 Years of School**	0 - No 1 - Yes

[a]CPS cases only.

*Included in both the short and long match lists.

**Included only in the short match list.

TABLE III.13

SUMMARY OF LONG LIST CELL MATCHING RESULTS
FOR THE SUPPORTED WORK AFDC EXPERIMENTAL SAMPLE

Cell	Number	Supported Work AFDC Sample		CPS AFDC Long List Matched Sample	
		Number	Percent	Number	Percent
1.	11009	19	2.4	9	1.6
2.	11109	22	2.8	7	1.3
3.	12000	10	1.3	177	31.9
4.	12001	28	3.5	97	17.5
5.	12011	4	0.5	17	3.1
6.	12100	4	0.5	71	12.8
7.	12101	33	4.1	58	10.5
8.	12111	9	1.1	8	1.4
9.	13099	50	6.3	11	2.0
10.	13100	13	1.6	5	0.9
11.	13101	76	9.5	4	0.7
12.	13119	23	2.9	16	2.9
13.	21000	14	1.8	9	1.6
14.	21001	45	5.6	5	0.9
15.	21019	11	1.4	4	0.7
16.	21100	12	1.5	5	0.9
17.	21101	62	7.8	5	0.9
18.	21119	14	1.8	6	1.1
19.	23099	53	6.6	4	0.7
20.	23199	80	10.0	5	0.9
21.	31009	56	7.0	4	0.7
22.	31019	18	2.3	5	0.9
23.	31109	81	10.1	7	1.3
24.	31119	28	3.5	6	1.1
25.	33999	34	4.3	9	1.6
Total		799	100.0	554	100.0

NOTE: Initially, all Supported Work and CPS sample members were assigned to cells defined by unique combinations of the first five characteristics shown in Table III.12. Cells were then collapsed so as to achieve cells with a minimum of four cases from each sample. Collapsing occurred across characteristics in the inverse order of their presentation in Table III.12. A value of "9" in a cell number indicates that cells were collapsed on the attribute denoted by that digit.

TABLE III.14

NUMBER OF CASES IN COLLAPSED CELLS FROM THE SUPPORTED WORK AFDC SAMPLE AND A MATCHED CPS SAMPLE, USING THE LONG LIST OF MATCH VARIABLES

Cell Number	Supported Work AFDC Sample		CPS AFDC Long List Matched Sample	
	Exact Match	In Collapsed Cells	Exact Match	In Collapsed Cells
Earnings, Pre-Base Year	799	0	554	0
Change in SSA Earnings, Pre-Base Year to Base Year	799	0	554	0
Age < 35	765	34	545	9
> 0 Weeks of Employment in Base Year	582	217	525	29
Black, Non-Hispanic	310	489	461	93

NOTE: The CPS sample used was the AFDC Stage 2 sample described in Table III.3.

Work experimental and CPS Stage 2 cases across the three active AFDC short-list match cells are presented in Table III.15. Note that 100 percent of the available Supported Work and CPS cases were matched.

As shown in Table III.16, the CPS AFDC long-list matched sample closely resembles Supported Work experimentals and controls with respect to base-year SSA earnings, change in SSA earnings, and age. This resemblance is greater than that achieved by either the short-list matched sample or the two random comparison samples. The short-list matched sample resembles the Supported Work samples more closely than the long-list matched sample with respect to race and education. In general, both matched CPS AFDC samples are more comparable to the Supported Work AFDC samples than are the CPS AFDC random comparison samples (Table III.4).

The mean earnings of the CPS AFDC long-list matched sample for the period 1972-1979 closely resemble those of the Supported Work AFDC control group. This close correspondence and the substantial improvement achieved relative to the Stage 2 random comparison sample can be observed by comparing Figure III.4 with Figure III.2. In 1977, a year of large in-program earnings for Supported Work participants, the mean earnings of the long-list matched sample differed from those of the Supported Work control group by only $225 (Appendix Table B.2). A simple difference-in-means estimate of the impact of the Supported Work program on 1977 earnings using the long-list matched sample differs only by 16 percent from the estimate obtained using the Supported Work control sample. Similarly encouraging results were obtained for the primarily post-program years of 1978 and 1979.

TABLE III.15

SUMMARY OF SHORT LIST CELL MATCHING RESULTS FOR THE SUPPORTED WORK AFDC EXPERIMENTAL SAMPLE

Cell Number[a]	Supported Work AFDC Sample		CPS AFDC Short List Matched Sample	
	Number	Percent	Number	Percent
1. 00	130	16.3	346	38.1
2. 01	643	80.5	223	24.5
3. 19[b]	26	3.3	340	37.4
Total	799	100.0	909	100.0

[a]The first digit of the cell number indicates categories of the variable "$\geq$ 12 Years of School," and the second digit indicates categories of the variable "Black, Non-Hispanic." Codes for the categories of these variables are provided in Table III.12.

[b]To avoid having fewer than 4 cases in a cell, we combined cases with 12 or more years of school into a single cell collapsed over racial categories.

TABLE III.16

PERCENT DISTRIBUTIONS OF CASES FROM THE SUPPORTED WORK AFDC SAMPLES AND WEIGHTED DISTRIBUTIONS OF MATCHED CPS AFDC CASES FROM THE SHORT LIST MATCH AND THE LONG LIST MATCH

Characteristic	Supported Work		CPS Comparison Groups	
	Experimentals	Controls	Short List	Long List
Age				
Younger than 21	1.0	0.5	3.7	5.7
21 - 34	58.1	59.0	32.5	52.5
35 or Older	40.8	40.0	63.8	41.9
Missing	0.1	0.5	0.0	0.0
Average	**33.3**	**33.5**	**38.5**	**33.3**
Race/Ethnicity				
Black, Non-Hispanic	83.5	81.6	81.6	52.7
Hispanic	12.0	13.0	4.4	12.3
White/Other	4.5	5.5	14.0	35.0
Education				
Fewer than 12 Years	69.7	69.2	69.2	62.9
12 or More Years	30.3	30.8	30.8	37.1
Missing	0.0	0.0	0.0	0.0
Average Years	**10.3**	**10.3**	**10.3**	**9.7**
Occupation, Baseline Interview				
Sales/Clerical	7.3	8.0	5.5	9.2
Service	10.3	8.4	16.7	19.1
Professional	1.2	1.6	2.0	2.7
Other	6.7	7.7	7.9	4.9
Missing/None	74.5	74.3	67.9	64.1
Public Sector Employment, Baseline Interview				
Yes	N.A.	N.A.	3.4	7.2
No			96.6	92.8
Percent of Time Employed, Base Year				
0	80.2	78.2	67.9	64.1
1 - 49	15.1	16.8	11.8	17.1
50 - 89	2.9	3.4	10.9	9.3
$\geq$ 90	1.4	0.7	9.4	9.5
Missing	0.4	0.9	0.0	0.0

TABLE III.16 (continued)

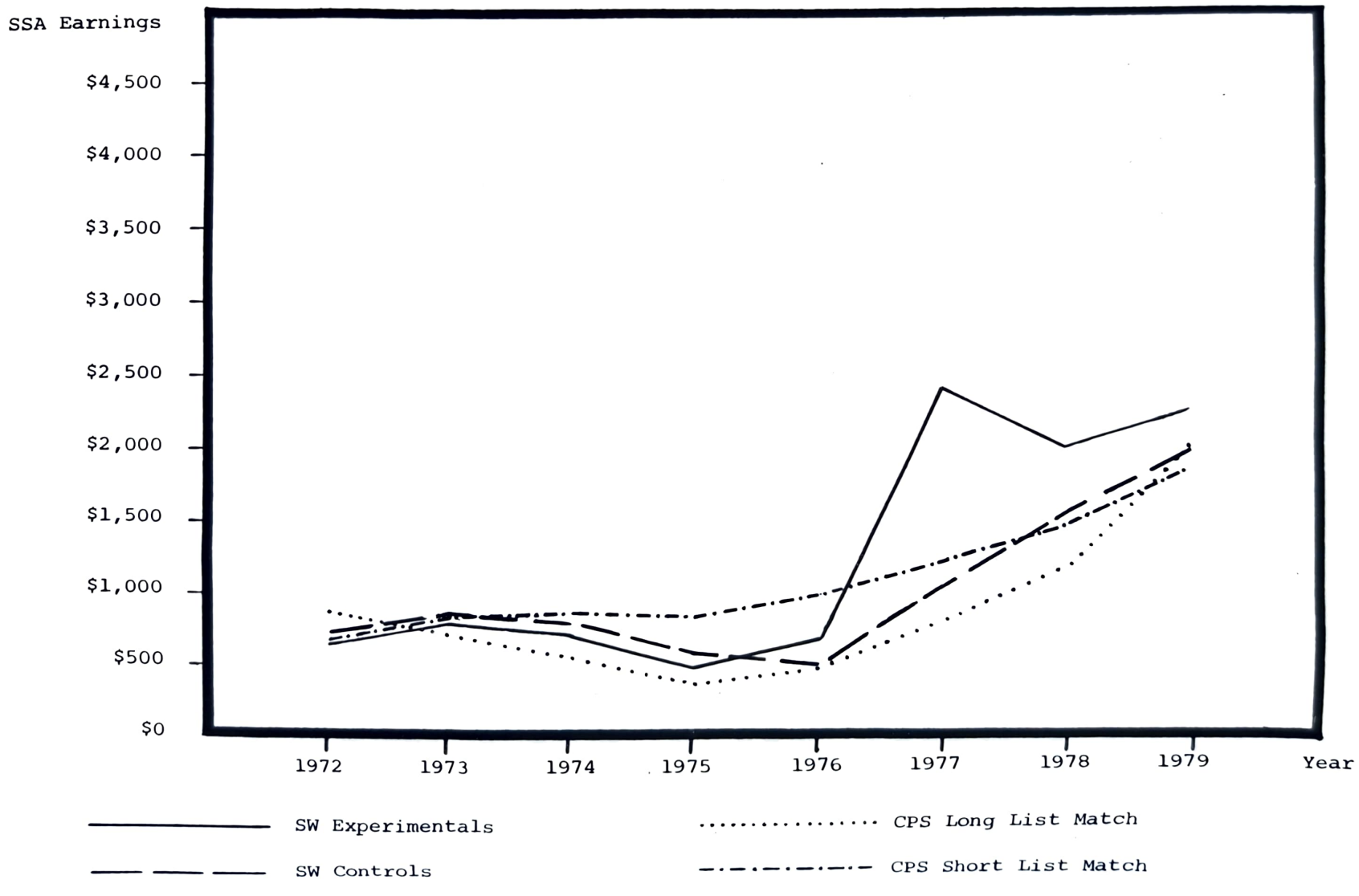
Characteristic	Supported Work		CPS Comparison Groups	
	Experimentals	Controls	Short List	Long List
SSA Earnings, Base Year[a]				
$0	10.6	5.4	65.6	34.0
$1 - $1,000	79.4	84.0	11.7	55.2
$1,001 - $5,000	10.0	10.6	19.9	10.7
≥ $5,000	0.0	0.0	2.8	0.0
Average	**$459**	**$463**	**$805**	**$352**
Change in SSA Earnings, Pre-Base Year to Base Year				
Negative	47.9	56.2	19.2	47.8
Zero[b]	16.2	12.3	59.3	11.5
Positive	35.9	31.4	21.5	40.7
Average Change	**$-55**	**$-153**	**$-20**	**$-86**
Welfare Receipt, Base Year				
Yes	99.9	99.2	100.0	100.0
No	0.0	0.3	0.0	0.0
Missing	0.1	0.5	0.0	0.0
Average Monthly Family Income, Base Year	**$310**	**$309**	**$566**	**$530**
Age of Youngest Child, Baseline Interview				
Younger than 6	2.1	2.6	0.0	0.0
6 - 12 Years	71.1	72.3	62.8	68.0
12 or Older	25.6	23.9	37.2	32.0
Missing	1.2	1.8	0.0	0.0
Number of Cases	800	802	909	554

[a]SSA earnings for Supported Work cases are group mean values rather than values for individuals. SSA earnings for CPS cases are values for individuals.

[b]The zero change in the SSA earnings category includes cases whose SSA earnings did not change at all, 5 percent of the smallest negative-change cases, and 5 percent of the smallest positive-change cases.

FIGURE III.4

AVERAGE SSA EARNINGS FOR AFDC WOMEN: SUPPORTED WORK EXPERIMENTALS AND CONTROLS, AND WEIGHTED CPS SHORT AND LONG LIST MATCHES



Earnings correspondence with Supported Work controls was generally smaller for the short-list match group than for the long-list match group. However, as indicated by Figures III.2 and III.4, the short-list match produced a comparison group whose 1972-1979 earnings were more similar to those of the control group than was the case for the Stage 2 random sample. Of particular importance are the positive simple estimates of longer-run (1978 and 1979) program impacts obtained using the short-list match group, which contrast with the zero estimates obtained by using the random sample. The former compare favorably with results obtained using the Supported Work control group.

On the basis of earnings patterns, the CPS AFDC matched samples are better comparison groups than are the CPS youth matched samples for two reasons:

- The matched CPS AFDC cases were prescreened for base-year welfare receipt. No similarly discriminating screen was available for CPS youth.
- Recent past earnings are better predictors of future earnings for prime-age adults than for young adults. The average base-year age of women in the CPS AFDC long-list matched sample was 33, compared with 18 for individuals in the CPS youth long-list matched samples.

D. STATISTICAL MATCHING

Statistical matching is the third procedure we used to draw comparison groups for Supported Work experimental samples from the 1976 and 1977 CPS general screen file. Values of 1979 SSA earnings were imputed to Supported Work experimental and CPS Stage 2 cases on the basis of their baseline characteristics and earnings histories. Each Supported Work case was matched with the CPS case that had the most similar imputed earnings.

Because CPS cases were selected only once, it was unnecessary to compute analysis weights for them.

This approach to statistical matching was designed to produce a comparison file with actual values of the match variable (1979 earnings in this application) that were similar to those which experimental cases would have had had they received no treatments (i.e., they are similar to those of the true control cases). However, the demographic and other characteristics of matched comparison cases may differ substantially from those of experimentals. When such differences are large, the reliability of inferences based on analyses of the statistically matched experimental and comparison samples should be questioned.

1. Statistical Matching for Youth

The ideal sample to use as the basis for estimating the parameters in an earning model is the Supported Work control group. However, in cases where comparison-group methodologies are relevant and control groups are unavailable, so we conducted our statistical matching based on coefficients from a 1979 earnings equation estimated for the CPS youth Stage 2 sample (see Appendix Table B.3). Our alternative was to use the Supported Work experimental group to estimate the earnings equation. However, in view of the fact that an F-test conclusively rejected the hypothesis that the regression coefficients are equal for the CPS and the experimental samples,

it seemed best to rely on the parameter estimates that were uncontaminated by program impacts--those from the CPS sample.[1]

Distributions of actual and the resulting imputed 1979 earnings are presented in Table III.17 for Supported Work youth experimentals and CPS Stage 2 youth. Three features of these distributions should be noted:

1. A large proportion of CPS youth had no SSA earnings in 1979. These represent observations on a truncated (at zero) earnings distribution. Under these conditions, OLS estimates of parameters in the CPS earnings equation are biased. Tobit analysis could have been used to obtain unbiased estimates; however, time limitations necessitated using OLS.

2. Predicted 1979 earnings tend to exceed actual 1979 earnings. This is true for 77 percent of the Supported Work youth and for 59 percent of the CPS youth. The above-noted bias in the OLS estimates of the earnings equation may account for some of this systematic error.

3. In every category of predicted earnings, the number of CPS cases exceeds the number of Supported Work cases. Thus, good potential matches on predicted earnings exist for all Supported Work youth.

In part, these problems with the statistical match can arise from our use of an OLS regression model as the basis for computing predicted earnings rather than the use of a Tobit model, which can properly account for the bounded nature of the dependent variable. In future work, it would be instructive to compute predicted values using models that address this problem. However, such procedures are expected to alter the distribution

---

[1]In fact, the failure to pass an F-test of the comparability of the earnings models for experimentals and the CPS sample calls into question the appropriateness of using the statistical matching procedure. However, a similar test has been failed in other studies that have used statistical matching to construct comparison groups, making it an interesting issue as to whether sound comparison samples can be identified in such cases (see for example, Dickinson et al., 1984).

TABLE III.17

DISTRIBUTIONS OF ACTUAL 1979 SSA EARNINGS AND 1979 SSA EARNINGS AS PREDICTED BY THE STATISTICAL MATCH EARNINGS EQUATION FOR THE SUPPORTED WORK YOUTH EXPERIMENTAL AND CPS STAGE 2 YOUTH SAMPLES

1979 SSA Earnings	Supported Work Youth				CPS Stage 2 Youth					
	Actual		Predicted		Actual				Predicted (Ungrouped Data)	
					Ungrouped Data		Grouped Data			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
≤ $0	0	0.0	5	0.9	695	29.3	0	0.0	25	1.1
$1 – $1,000	122	21.8	22	3.9	171	7.2	29	9.0	163	6.9
$1,001 – $2,000	163	29.1	54	9.6	196	8.3	35	10.9	297	12.5
$2,001 – $3,000	128	22.9	137	24.5	171	7.2	51	15.9	347	14.7
$3,001 – $4,000	105	18.8	236	42.1	147	6.2	47	14.6	363	15.3
$4,001 – $5,000	26	4.6	79	14.1	128	5.4	50	15.6	361	15.2
≥ $5,001	16	2.9	27	4.8	860	36.3	109	34.0	812	34.3
Total	560	100.0	560	100.0	2368	100.0	321	100.0	2368	100.0

NOTE: Predicted values of 1979 SSA earnings were generated by a single regression equation tht was estimated on the basis of ungrouped data for the full CPS Stage 2 Youth sample. This equation is shown in Table B.3.

primarily at the extremes and will affect the CPS and Supported Work Cases in a comparable way and so may have little overall affect on the statistical match results.

Acceptable matches were found for all but 2 of the 560 Supported Work youth for whom earnings could be imputed. The unmatched cases contained no CPS Stage 2 youth counterparts with imputed 1979 SSA earnings within $500 of their own. Characteristics of statistically matched Supported Work and CPS cases are given in Table III.18 and their mean earnings profiles are shwon in Figure III.5 The sample of statistically matched Supported Work youth consists primarily of black men, while the sample of statistically matched CPS youth consists primarily of white women. This outcome exemplifies a problem associated with this variant of statistical matching that was mentioned above: cases that match closely with respect to imputed earnings may differ substantially in the characteristics that determine earnings.

These substantial differences in the characteristics of the statistically matched experimental and comparison samples raise doubts about the reliability of program impact estimates based upon them. In response to these concerns, an alternate statistical match for youth was conducted. We again matched on predicted 1979 SSA earnings, but agreement on race and gender was also required. This was accomplished by conducting separate matches for each of six race/gender groups. Because of the small number of CPS Stage 2 cases relative to Supported Work experimental cases in several of these cells, we chose to match CPS cases with replacement. The matching algorithm gave preference to CPS cases from central cities and rejected matches for which the difference in predicted earnings exceeded $500. Full

TABLE III.18

PERCENT DISTRIBUTIONS OF CASES FROM THE SUPPORTED WORK YOUTH SAMPLES AND CPS YOUTH STATISTICAL MATCH SAMPLE

Characteristic	Supported Work		CPS Statistical Match Sample
	Matched Experimentals	Controls	
Gender			
Male	88.7	86.9	28.7
Female	11.3	13.1	71.3
Age			
Younger than 19	61.3	56.6	64.9
19 or Older	37.1	41.9	35.1
Missing	1.6	1.5	0.0
Average	**17.8**	**18.0**	**17.9**
Race/Ethnicity			
Black, Non-Hispanic	72.6	69.9	14.2
Hispanic	17.6	19.6	11.5
White/Other	8.1	8.6	74.4
Missing	1.8	1.9	0.0
Education			
Fewer than 12 Years	99.3	99.4	100.0
12 or More Years	0.7	0.6	0.0
Missing	0.0	0.0	0.0
Average Years	**9.3**	**9.7**	**9.7**
Marital Status, Baseline Interview			
Married	4.5	3.5	27.2
Not Married	95.5	96.3	72.8
Missing	0.0	0.2	0.0
Occupation, Baseline Interview			
Sales/Clerical	10.4	11.2	16.5
Service	25.6	21.7	31.7
Professional	5.6	4.0	1.6
Other	27.6	29.5	21.0
Missing/None	30.8	33.6	29.2
Public Sector Employment, Baseline Interview			
Yes	n.a.	n.a.	4.7
No			95.3

TABLE III.18 (continued)

Characteristic	Supported Work: Matched Experimentals	Supported Work: Controls	CPS Statistical Match Sample
Percent of Time Employed, Base Year			
0	99.1	43.7	89.1
1 - 49	0.9	40.4	10.9
50 - 89	0.0	12.1	0.0
≥ 90	0.0	1.8	0.0
Missing	0.0	2.1	0.0
SSA Earnings, Base Year[a]			
$0	1.4	0.0	31.2
$1 - $1,000	84.4	87.6	38.5
$1,001 - $5,000	14.2	12.4	30.3
> $5,000	0.0	0.0	0.0
Average	**$580**	**$550**	**$751**
Change in SSA Earnings, Pre-Base Year to Base Year[a]			
Negative	29.0	34.2	22.4
Zero[b]	10.4	12.2	25.8
Positive	60.6	53.5	51.8
Average Change	**$127**	**$98**	**$269**
Welfare Receipt, Base Year			
Yes	24.6	24.6	6.1
No	74.7	73.8	93.9
Missing	0.7	1.6	0.0
Average Monthly Family Income--Baseline Year	**$92**	**$95**	**$277**
Number of Cases	558	678	558

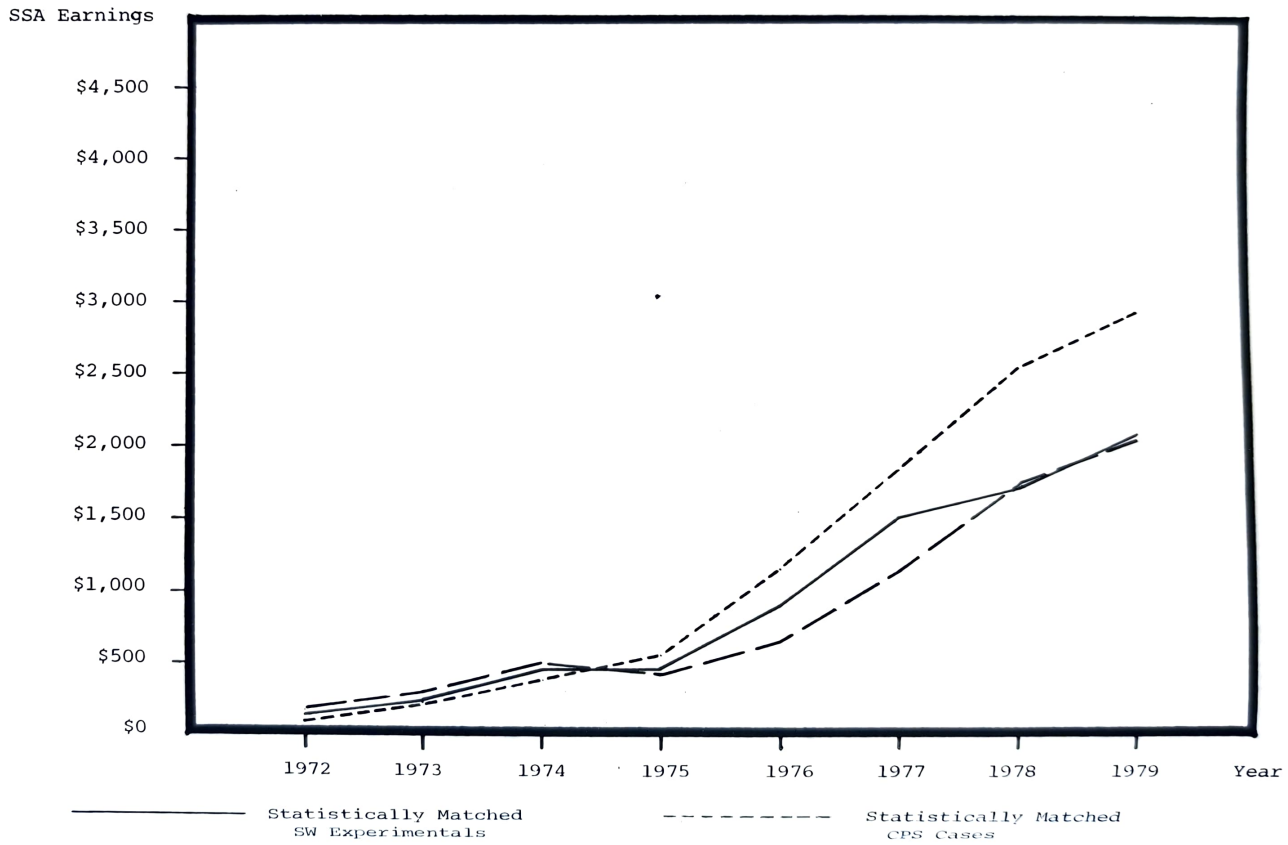
[a]SSA earnings for Supported Work cases are group mean values rather than values for individuals. SSA earnings for CPS cases are values for individuals.

[b]The zero change in the SSA earnings category includes cases whose SSA earnings did not change at all, 5 percent of the smallest negative-change cases, and 5 percent of the smallest positive-change cases.

n.a. means not available.

FIGURE III.5

AVERAGE SSA EARNINGS FOR YOUTH: SUPPORTED WORK CONTROLS AND STATISTICALLY MATCHED CASES FROM THE SUPPORTED WORK EXPERIMENTAL SAMPLE AND CPS RANDOM SAMPLE 2



details on the alternate statistical match, including mean earnings profiles and a table describing the characteristics of the matched samples, are provided in Appendix E.

As referenced in Chapter IV, our evaluation of alternative comparison groups provided no evidence that the alternative statistical matching procedure is better than the initial procedure described above. This conclusion is based upon a comparison of estimates of the Supported Work program's net impact on earnings obtained on the basis of the true control group with estimates obtained on the basis of the initial and alternate statistically matched comparison groups.

2. Statistical Matching for AFDC Recipients

As with youth, a 1979 SSA earnings equation estimated on the basis of CPS data was used to impute earnings to Supported Work AFDC experimental and CPS Stage 2 AFDC cases. This equation is shown in Appendix Table B.4. Each of the 799 Supported Work cases for which earnings could be imputed was matched to the CPS case which had the most similar imputed earnings. Supported Work cases whose imputed earnings differed from those of their CPS matches by more than $500 were excluded from the final file of 582 matched pairs of experimental and comparison cases. As with youth, the CPS cases in this file differ substantially from the Supported Work cases with respect to certain key personal characteristics, suggesting that the matching procedure should be revised so as to ensure acceptable levels of conformity between the matched samples for critical variables.

Table III.19 shows a concentration of CPS AFDC Stage 2 cases at zero 1979 SSA earnings, suggesting that, as with youth, OLS estimates of the earnings equation may be biased. More importantly, this table reveals

TABLE III.19

DISTRIBUTIONS OF ACTUAL 1979 SSA EARNINGS AND 1979 SSA EARNINGS AS PREDICTED BY THE STATISTICAL MATCH EARNINGS EQUATION FOR THE SUPPORTED WORK EXPERIMENTAL AND CPS STAGE 2 AFDC SAMPLES

1979 SSA Earnings	Supported Work AFDC				CPS Stage 2 AFDC					
	Actual		Predicted		Actual				Predicted (Ungrouped Data)	
					Ungrouped Data		Grouped Data			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
≤ $0	0	0.0	1.0	1.3	504	55.4	7	6.0	18	2.0
$1 - $1,000	164	20.5	147	18.4	42	4.6	30	25.6	247	27.2
$1,001 - $2,000	192	24.0	405	50.7	66	7.3	29	24.8	305	33.6
$2,001 - $3,000	212	26.5	219	27.4	48	5.3	12	10.3	110	12.1
$3,001 - $4,000	134	16.8	18	2.3	39	4.3	17	14.5	74	8.1
$4,001 - $5,000	50	6.3	0	0.0	28	3.1	12	10.3	45	5.0
≥ $5,001	47	5.9	0	0.0	182	20.0	10	8.5	110	12.1
Total	799	100.0	799	100.0	909	100.0	117	100.0	909	100.0

NOTE: Predicted values of 1979 SSA earnings were generated by a single regression equation that was estimated on the basis of ungrouped data for the full CPS Stage 2 AFDC sample. This equation is sown in Table B.4.

a shortage of good potential CPS matches for Supported Work cases with imputed earnings in the $1,001-$3,000 range. The shortage is attributable to the smallness of the CPS Stage 2 AFDC sample and the failure of its relative frequency distribution to coincide closely with that of the AFDC experimentals. These factors prevented us from obtaining an acceptably close (within $500) match on imputed earnings for over 200 experimental cases.

The distributions in Table III.20 reveal that the sample of statistically matched Supported Work cases consists primarily of black women and women who were younger than 35 in the base year. In contrast, the sample of statistically matched CPS cases consists primarily of white women and women who were older than 34 in the base year. While the imputed 1979 SSA earnings of these two matched samples of women are similar, we are concerned that the CPS sample is not likely to constitute a reliable comparison group for assessing the net impact of the Supported Work program on earnings. As indicated by Figure III.6, the mean earnings of this comparison group were below those of Supported Work controls throughout the 1972-1979 period. The difference was greatest during the critical post-program years of 1978 and 1979.

The alternate statistical matching procedure was implemented on the Supported Work experimental and CPS Stage 2 random sample of AFDC women so as to obtain a matched sample with greater demographic correspondence. The implementation of this procedure and the characteristics of the resultant matched samples are documented in Appendix E. Separate matches were conducted on six different race/age groups, following the same alternate procedure as described above for youth. As noted in Chapter IV, estimates

TABLE III.20

PERCENT DISTRIBUTION OF CASES FROM THE SUPPORTED WORK AFDC SAMPLES AND CPS AFDC STATISTICAL MATCH SAMPLE

Characteristic	Supported Work		CPS Statistical Match Sample
	Matched Experimentals	Controls	
Age			
Younger than 21	0.9	0.5	3.9
21 - 34	56.0	59.0	36.6
35 or Older	43.1	40.0	59.5
Missing	0.0	0.5	0.0
Average	**33.8**	**33.5**	**37.6**
Race/Ethnicity			
Black, Non-Hispanic	84.5	81.6	37.8
Hispanic	11.7	13.0	13.2
White/Other	3.8	5.5	49.0
Education			
Fewer than 12 Years	69.1	69.2	68.9
12 or More Years	30.9	30.8	31.1
Missing	0.0	0.0	0.0
Average Years	**10.3**	**10.3**	**9.4**
Occupation, Baseline Interview			
Sales/Clerical	6.0	8.0	5.2
Service	12.5	8.4	15.0
Professional	1.5	1.6	1.0
Other	2.7	7.7	2.6
Missing/None	77.3	74.3	76.2
Public Sector Employment, Baseline Interview			
Yes	n.a.	n.a.	4.8
No			95.2
Percent of Time Employed, Base Year			
0	99.5	78.2	96.2
1 - 49	0.5	16.8	3.8
50 - 89	0.0	3.4	0.0
$\geq$ 90	0.0	0.7	0.0
Missing	0.0	0.9	0.0

TABLE III.20 (continued)

Characteristic	Supported Work: Matched Experimentals	Supported Work: Controls	CPS Statistical Match Sample
SSA Earnings, Base Year[a]			
$0	11.5	5.4	76.8
$1 - $1,000	81.8	84.0	16.2
$1,001 - $5,000	6.7	10.6	7.0
≥ $5,000	0.0	0.0	0.0
Average	**$407**	**$463**	**$173**
Change in SSA Earnings, Pre-Base Year to Base Year[a]			
Negative	47.9	56.2	21.6
Zero[b]	17.7	12.3	64.8
Positive	34.4	31.4	13.6
Average Change	**$-55**	**$-153**	**$-203**
Welfare Receipt--Base Year			
Yes	100.0	99.2	100.0
No	0.0	0.3	0.0
Missing	0.1	0.5	0.0
Average Monthly Family Income, Base Year	**$307**	**$309**	**$567**
Age of Youngest Child--Baseline Interview			
Younger than 6	1.7	2.6	0.0
6 - 12 Years	71.5	72.3	67.0
12 or Older	25.8	23.9	33.0
Missing	0.0	1.8	0.0
Number of Cases	582	802	582

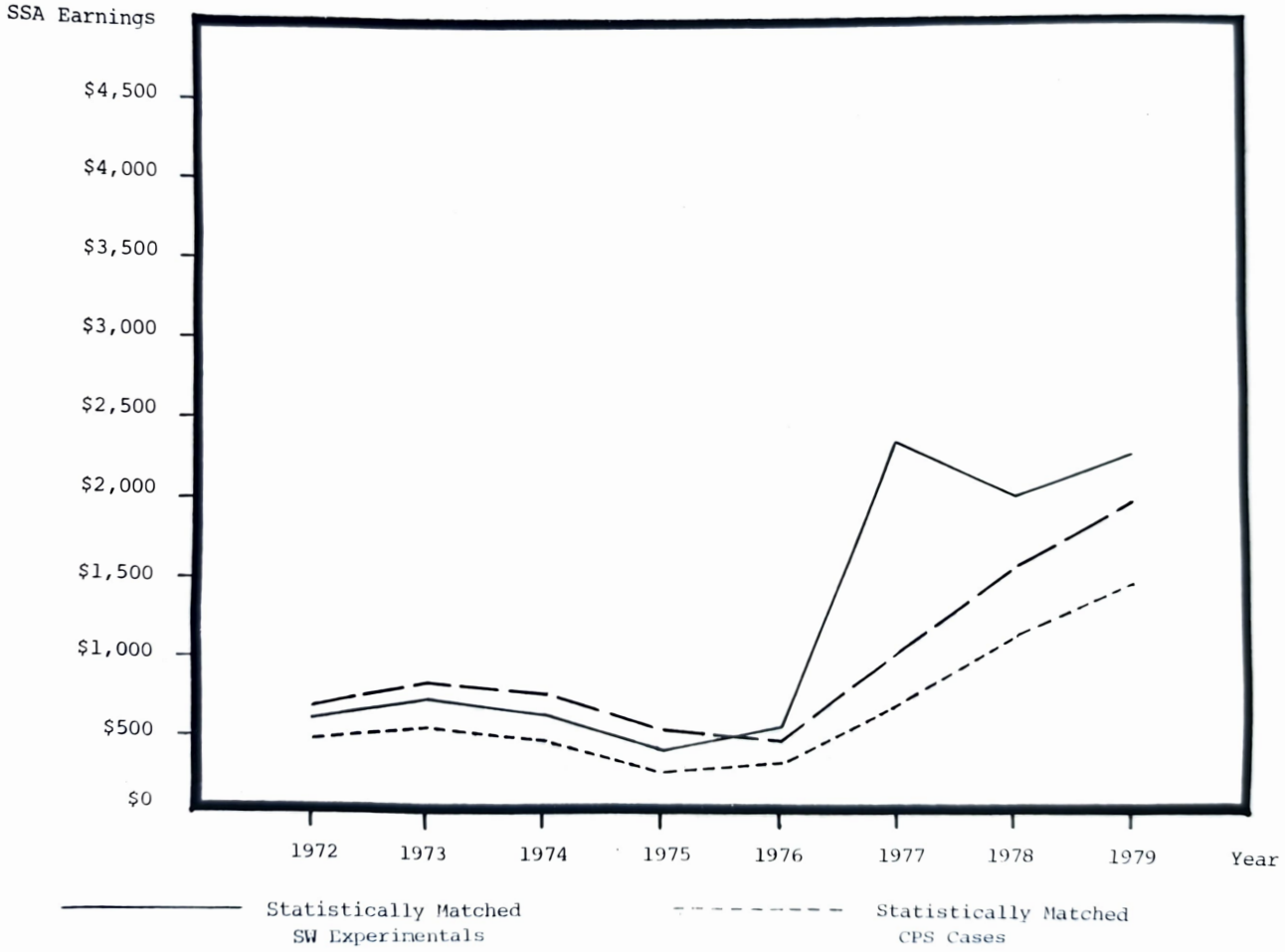
[a]SSA earnings for Supported Work cases are group mean values rather than values for individuals. SSA earnings for CPS cases are values for individuals.

[b]The zero change in the SSA earnings category includes cases whose SSA earnings did not change at all, 5 percent of the smallest negative-change cases, and 5 percent of the smallest positive-change cases.

n.a. means not available.

FIGURE III.6

AVERAGE SSA EARNINGS FOR AFDC WOMEN: SUPPORTED WORK CONTROLS AND STATISTICALLY MATCHED CASES FROM THE SUPPORTED WORK EXPERIMENTAL SAMPLE AND CPS RANDOM SAMPLE 2



of program net impacts based on the alternate statistical match comparison sample differ little from those based on the initial statistical match comparison sample.

IV. EVALUATION OF ALTERNATIVE COMPARISON GROUPS

The overall net-impact results from the Supported Work demonstration are generally consistent with the findings from numerous other evaluations of program interventions for youth and AFDC recipients--the program had significant and lasting positive impacts on the earnings of AFDC recipients and only short-run (in-program) earnings effects on youth (see Hollister et al., 1984). Given that the Supported Work evaluation was based on an experimental design and that longitudinal data collection was undertaken for relatively large samples, it is unlikely that the evaluation results fail to reflect the true impacts of the program reasonably well. Our objective in this chapter is to assess the validity and reliability of the estimated net program impacts that would likely have been generated from an evaluation effort in which it was necessary to construct comparison groups from existing data. To supplement our general assessment of the overall quality of conclusions based on an evaluation design that relies on ex-post comparison-group methodologies, we explore a number of more detailed issues: (1) the general adequacy of existing data bases for drawing comparison samples for participants in special employment programs, (2) the importance of the methodology used to select the comparison group (e.g., random samples, cell-matched samples, or statistical-match samples); (3) the sensitivity of the net-impact results to the analytic model used; and (4) the effects of using weighted versus unweighted data.

In this chapter, we first summarize the results of our efforts to construct comparison groups, and offer some generalizations about the adequacy of existing data in doing so. In the second section, we

outline the general analytic approach adopted to assess the comparison-group evaluation methodologies and present the results of this methodological study. In the final section, we discuss the findings of this study within the context of the net-impact findings from studies that have relied on CPS and CLMS data to assess the net impacts of CETA.

A. PRACTICAL CONSIDERATIONS IN ADOPTING COMPARISON GROUP METHODOLOGIES

The purpose of using a comparison group in a net-impact evaluation is to address the following question: "What would the behavior of and outcomes for the participant group have been had the program not existed or had they been denied participation?" Several conditions must be met to achieve this objective. First, the comparison group must be either similar to program participants along all dimensions that affect the targeted or intended program outcome or different only in ways that have been measured. Second, the program-participant and comparison-group members must exhibit a sufficient degree of similarity so that a common analytic model can be applied to both groups. Third, the comparison group must be large enough to generate reasonably precise estimates of the outcome measures. In constructing comparison groups for the Supported Work youth and AFDC samples from the CPS, we found potential problems with each of these conditions.

1. Comparability of Participant and Comparison Group Characteristics

Not surprisingly, the less narrowly defined random samples of youth and AFDC recipients from the CPS (Stage 1 samples) do not match the program-participant samples well along most measurable characteristics. First, the Supported Work youth, all of whom were school dropouts and who

came primarily from large urban areas, obviously are not representative of the population of 16- to 20-year-olds; second, the Supported Work AFDC recipients, none of whom had a child younger than age 6, differ substantially from the 54 percent of AFDC recipients who do have young children.

Although the comparability of the participant and comparison samples was substantially improved by limiting the comparison samples to CPS observations that met the more narrowly defined Supported Work youth and AFDC target populations (Stage 2 samples), large differences persisted. Most notably, minorities are greatly underrepresented in both the youth and AFDC comparison groups, undoubtedly due to the urban focus of the Supported Work program[1]; among the youth comparison group, both females and married youth are overrepresented, since Supported Work attracted predominantly unmarried minority males. The CPS sample of AFDC recipients with no young children differs from the Supported Work AFDC sample with respect to its higher average age and higher rate of employment in the base year.

The statistical-match comparison samples for both the youth and AFDC recipients exhibit characteristics that, with a few exceptions, are similar to those of the Stage 2 random samples of CPS cases. To a lesser extent, this is also true of the cell-match samples, primarily due to the relative paucity of CPS observations that mirror the unique combinations of characteristics of many of the Supported Work participants. This point is illustrated by the frequency with which racial distinctions need be ignored

---

[1] The CPS sample was too small to permit us to restrict the comparison sample to urban residents.

in obtaining CPS matches for Supported Work cases in the cell-matching procedure, and by the large range of values for the match weights (0.31 to 10.00 for youth and 0.56 to 19.00 for AFDC women). The initial statistical matching algorithm used did not impose penalties for differences between the participant observation and the match case on particular characteristics (see Chapter III), and thus the characteristics of the statistical match comparison samples differ from their participant group counterparts in some key dimensions. However, given the overall sample-size constraints, when we did match within strata, we tended to achieve better matches only at the cost of using some CPS cases multiple times.[1]

2. The Similarity of Behavioral Models for Participant and Comparison Group Members

In view of the pervasive differences in the characteristics of the participant and comparison samples and the observed differences in the earnings trends over time (see Chapter III), it is important that we consider whether or not the underlying relationships are similar between the measurable characteristics and earnings (or the other outcomes of interest) of the comparison groups and the groups they were chosen to represent (i.e., participants, had they not had an opportunity to enroll in Supported Work). If not, analytic models that pool the participant group with a comparison group would be inappropriately specified and could be expected to yield biased estimates of program impacts.

The standard Chow test provides a relatively weak test of this similarity between the participant and comparison groups. Thus, we

---

[1]Details on statistical matching within strata are provided in Appendix E.

performed such tests using a basic earnings model (discussed further in Section B). The results of these tests, summarized in Table IV.1, suggest that once the basic demographic and background characteristics are controlled for, many of the comparison groups are sufficiently similar to the experimental group that there is no objective basis for ruling out conducting the net-impact analysis. However, for the majority of the AFDC samples and for three of the youth samples, there is evidence of different underlying models for the participant and comparison samples.

3. Adequacy of the CPS Sample Size

For a number of reasons, few data bases are superior to the CPS for constructing comparison groups to evaluate employment programs. These include the CPS's large size, its relatively comprehensive data, the longitudinal element in the SSA earnings data that have been appended to the files, and the fact that it is representative of the entire population. However, the CPS does have limitations relative to its size and/or stratification and to its data elements.

Despite the large size of the CPS sample (over 200,000 observations per year), this sample includes relatively few young out-of-school youth and AFDC recipients. Combining the March 1976 and March 1977 CPS observations yields only 2,368 young school dropouts and 1,995 female AFDC recipients, fewer than half of whom have no young children (see, respectively, Tables III.1 and III.3). Furthermore, of these individuals, only 57 percent of the youth and 67 percent of the AFDC recipients in the CPS live in SMSAs, which is where the vast majority of the Supported Work program participants resided. Because of the relatively small proportion of the CPS cases residing in SMSAs, we could not restrict our comparison

TABLE IV.1

RESULTS OF CHOW TEST OF THE SIMILARITY OF EARNINGS MODELS BETWEEN THE EXPERIMENTAL AND VARIOUS COMPARISON SAMPLES (F-ratios)

Target Group/ Outcome Measure	Random Samples		Cell Match Samples		Statistical Match Sample
	Stage 1 Screens	Stage 2 Screens	Short Match List	Long Match List	
Youth					
1977 Earnings	1.53*	1.18	0.87	1.42	0.69
1979 Earnings	1.96**	1.17	0.85	4.57**	1.01
AFDC Recipients					
1977 Earnings	1.50	1.80*	2.12**	2.31**	1.21
1979 Earnings	1.93**	1.90**	1.51	2.26**	1.10
Number of Observations					
Youth					
Controls	87	87	87	87	87
Comparison	158	321	320	112	71
AFDC					
Controls	107	107	107	107	107
Comparisons	266	117	115	73	74

NOTE: These tests were performed by estimating separate models for the Supported Work control samples and various comparison samples and models that pool the control observations with each of the comparison samples. The control sample was used since its earnings behavior was uncontaminated by the program's impacts.

*Significant at the 5 percent level.
**Significant at the 1 percent level.

groups only to urban residents--a limitation that could cause inconsistencies in the behavioral relationships that underlie the analytic models (see above).

Other limitations of the CPS pertain to the available data elements. In particular, the CPS does not include either geographic identifiers or indicators of program participation by sample members, both of which could be important sample screens or match variables. Furthermore, the unavailability of less aggregate earnings data than the annual Social Security data preculde both a close matching on enrollment data and on transitory declines in employment.

However, while the utility of the CPS is reduced somewhat by these limitations, the most prominent characteristics of the Supported Work participants considered in this study (i.e., school dropout for youth and AFDC receipt for women) are sufficiently similar to large portion of former CETA and current JTPA participants that it is still useful to assess the quality of comparison samples that can be constructed from it. Such generalizations of the results can be especially important in view of the ongoing design of the replacement data base for the CLMS (the JTLS)[1] and the renewed emphasis on comparison-group evaluation designs for cost reasons.

B. THE QUALITY OF NET IMPACT ESTIMATES GENERATED USING THE COMPARISON GROUPS

The central objectives of this study are (1) to determine whether a comparison-group design, rather than an experimental design, could have

---

[1]JTLS stands for the Job Training Longitudinal Survey.

been used to generate accurate conclusions about the success of the national Supported Work demonstration and (2) to assess the extent to which the conclusions regarding the Supported Work demonstration based on a comparison-group design can be generalized to other program-evaluation efforts, especially evaluations of CETA and JTPA. In this section, we first outline the basic approach we have adopted to assess the reliability of net-impact estimates. We then summarize the results of this investigation. Section C discusses the generalizability of our conclusions for the Supported Work demonstration to other program evaluations.

1. Analytic Approach

Our approach for judging the overall quality of the net-impact estimates generated from the comparison groups is based on a comparison of these results with those generated from the true Supported Work control group. For each of the youth and AFDC groups, we have estimated net impacts on earnings for each of three calendar years (1977, 1978, and 1979)[1] by using three types of models.

Earnings Models. The first model is a simple comparison of the earnings gains between the base (pre-program) year and the outcome year for the participant group and the control or comparison group. This analytic model can be expressed as:

$$(1) \qquad (Y_i - Y_{0i}) = \alpha_1 + \beta_1 P_i + \varepsilon_i,$$

---

[1] As has been pointed out by Dickenson et al. (1984), the 1979 Social Security earnings data for the CPS sample include some values of zero that should be positive. For this reason, earnings of the comparison sample will be biased downward causing a positive bias in net program impact estimates.

where

Y denotes annual SSA earnings in the specified outcome year,

Y_0 denotes annual SSA earnings in the base year,

P is a binary variable denoting participant or control/comparison-group status,

ε is a random error term,

α and β are the coefficients to be estimated, and

i denotes the observation.

In this case, the impact of the program on annual earnings is measured by the estimate of β_1.

The second model is a basic earnings equation of the following form:

$$(2) \qquad Y_i = \alpha_2 + \beta_2 P_i + Z_i \delta + e_i,$$

where

Z represents a vector of control variables measuring factors believed to affect earnings,

δ is a vector of coefficients on the control variables,

and all other variables are defined similarly to their counterparts in equation (1). Again, the net program impacts in this model are measured by the estimated coefficient on the binary participation-status variable. This model differs from the first in that it attempts to control statistically for differences between the participant and comparison group along dimensions that may affect their earnings trends over time. Two critical assumptions underlie this model: (1) that the vector of control variables

fully accounts for factors that both pertain to program participation and affect the outcome of interest, and (2) that the underlying behavioral models of the determinants of earnings are similar for the participant and comparison groups. This equation is perhaps the most common form of the analytic model used to estimate net program impacts. It is sometimes applied as a cross-sectional model and estimated by ordinary least squares techniques and sometimes expanded to include (1) longitudinal observations on the sample members and (2) the corresponding time-specific and individual-specific error terms, and estimated by generalized least squares techniques. For our purposes, we have adopted the simpler cross-sectional model, for which we sacrifice only some efficiency in our estimates.

The third type of analytic model is a fixed-effect estimator, which can be expressed as follows:

$$(3) \qquad (Y_{ti} - Y_{0i}) = \alpha_3 + P^*_{ti}\beta_3 + (Z_{ti} - Z_{0i})\delta_3 + \varepsilon^*_{ti},$$

where

P^*_i is a vector of time and program status indicators,

Z_0 is a vector of base-period measures of the control variables,

Z_t is a vector of current-period measures of the control variables,

ε^*_{ti} has both a time-specific component $(\varepsilon_t - \varepsilon_0)$ and an individual-specific component $(\varepsilon_{ti} - \varepsilon_{0i})$.

This model essentially assumes that any unobserved and uncontrolled for factors that are correlated with program-participant status and earnings are constant over time, and, thus, that any bias in the net-impact estimates due to unobserved differences between the participant and comparison

groups can be controlled for by "differencing" a base-year and an outcome-year earnings equation.[1] The model also assumes that the earnings structures are similar for the participant- and comparison-group members, and that the basic models are similar across time periods. Especially with respect to youth, the latter condition may not hold.

In applying each of these analytic models, we have drawn on a fairly standard set of control variables for both the youth and AFDC samples. The ultimate choice of variables was based on several considerations: (1) the standard theory of labor supply, applied to these samples of young school dropouts and female welfare recipients, (2) consistency with the models used by Westat (1980) and proposed by Dickinson et al. (1983) in their analyses of CETA net impacts, to the extent that we found such consistency feasible and practical given the differences in the samples and the data bases being used, and (3) data constraints. Table IV.2 defines the control variables used in the basic versions of each of these three analytic models. Appendix Tables C.1 and C.2 present the means of these variables for the various samples.

Grouped Data. As we discussed in Chapter II, the net-impact analysis should be conducted using grouped rather than individual data, since the SSA earnings data for the Supported Work sample are mean values for groups of 7 to 10 individuals. Grouping the data was necessary to enable us to gain access to the social security data for this sample. Due to this limitation on the Supported Work data, we have ex-post grouped the

---

[1] A variant of this model might include proxy measures for the program participation decision. Also, it sometimes includes control variables that do not change over time and so would drop out of this formulation.

TABLE IV.2

DEFINITIONS OF CONTROL VARIABLES USED IN THE ANALYSIS

Mneumonic	Description	Analytic Model			Sample	
		Earning Gains	Basic Earning Equation	Fixed Effects Model[a]	Youth	AFDC Recipients
CPS76/77	Enrolled/interviewed March 1976 or Earlier		X	X	X	X
EXPERIM	Participant Status	X	X	X	X	X
AGE	Age in Years		X			X
AGESQ	Age in Years Squared		X	X		X
AGESCU	Age in Years Cubed			X		
TARGYTH	Younger than 19		X		X	
BLKFLG	Black, NonHispanic		X	X	X	X
HISFLG	Hispanic		X	X	X	X
ANYWORK	Any work, Base Year		X	X		X
WRKDUM2	Worked		X		X	
WRKDUM34	Worked				X	
WKSWORK	Weeks Worked, Base Year		X	X		X
WKSLOOK	Weeks Unemployed, Base Year		X	X	X	X
HOURS	Hours Worked Per Week, Base Year		X	X	X	X
SALECLER	Prior Occupation, Sales or Clerical		X	X	X	X
SERVICE	Prior Occupation-Service		X	X	X	X
PROFESS	Prior Occupation-Professional		X	X	X	X
MISSOCC	Prior Occupation-Code Missing		X	X	X	X
SSDELTA	Change in Earnings Pre Base Year to Base Year		X	X	X	X
SSPRI2	Earnings Pre Base Year		X	X	X	X
SSPRI3	Earnings 2 Years Prior to Base Year		X	X	X	X
PERCPYMI	Per Capita Family Income Base Year-Middle Third		X	X	X	X
PERCPYHI	Per Capita Family Income Base Year-Upper Third		X	X	X	X
EDDUM1	Less Than 10 years of Education		X	X	X	

TABLE IV.2 (Continuation)

Mneumonic	Description	Analytic Model			Sample	
		Earning Gains	Basic Earning Equation	Fixed Effects Model	Youth	AFDC Recipients
EDDUM2	10-11 Years of Education		X	X		X
AFDCRCP	AFDC Receipt Base Year		X		X	
MARFLAG	Married		X		X	
NEWSEX	Male		X		X	
LT12	No Dependent Younger Than 12		X	X		X
START	First Year any Sample Members were enrolled			X	X	X
MAIN	Main Enrollment Year for Sample			X	X	X
POST1	Main Enrollment Year Plus 1			X	X	X
POST2	Main Enrollment Year Plus 2			X	X	X
POST3	Main Enrollment Year Plus 3			X	X	X
T	Year (1 = 1974, ..., 6 = 1979)			X	X	X
TSQ	Year Squared (See Above)			X	X	X

[a]A more limited list of variables was included in one variant of the fixed effects model. This more restrictive set of variables included only those variables that change over time.

CPS comparison observations in a manner similar to the way in which the Supported Work cases were grouped (see Chapter II above), have constructed cell means for all analysis variables, and have conducted the net-impact analysis using the grouped data.

Doing so raises two analytic problems. The first pertains to the heteroskedasticity that occurs when using grouped data from cells of unequal sizes. This problem can be addressed quite simply by using weighted least squares, where all observations are weighted by the the square root of the number of cases in the cell.

The second problem pertains to the masking of limit observations in constructing the group means for the outcome measures. If micro-level data are used, the problem of limit observations can be dealt with quite simply by using maximum likelihood techniques such as double-bounded Tobit. However, the problem is analytically more complex when using grouped data, and the seriousness of the potential bias arising from this grouping is unknown. It would be useful in future work to test the sensitivity of the basic net impact results to adopting alternative procedures to account for this truncation.

Weighting. The concept underlying the cell-match strategy of comparison-group construction is as follows: if relatively small subgroups of participants who are identified by a unique combination of characteristics that are correlated with the outcome of interest can be matched with nonparticipants who exhibit the same combinations of characteristics, comparisons of the mean outcome measures for the participant- and comparison-group members in the cell will yield unbiased estimates of the program impacts for that particular subset of the population. The overall

effect would then be estimated as the mean of the outcome measure for all participant observations minus the mean for the comparison observations, where each comparison observation is weighted by the ratio of participant observations to comparison observations in its cell.

Within the context of net-impact estimates generated through regression analysis or other multivariate techniques, such weighting is necessary only insofar as the analytic model fails to fully control for factors that are correlated both with participant status and the outcome of interest. If this condition is met, results based on weighted and unweighted data should be comparable. However, to the extent that it is not met, using weighted data or controlling directly for cell assignment will generate impact estimates with more desirable statistical properties. For this reason, the main results reported herein are based on weighted data, and tests of the sensitivity of the results to weighting are presented separately. (See further discussion of weighting in Section 2 below and in Appendix F.)

Sensitivity Analyses. Having conducted the basic analysis using the models and procedures described above, we conducted a number of investigations to determine the robustness of the results. Among the sensitivity analyses are those which assess (1) the implications of pooling various subsets of the youth and AFDC samples, (2) the implications of including additional control variables and of altering the definitions of certain included variables, (3) the importance of weighting, and (4) the implications of using grouped as opposed to micro-level data. In all of this sensitivity analysis, the importance of various econometric assump-

tions with respect to the substantive policy conclusions that can be drawn from an evaluation has been emphasized.

2. Results of the Assessment of the Quality of the Net Impact Estimates

The results of this analysis demonstrate the sensitivity of the net-impact estimates both to the particular analytic models used with comparison-groups and to the methodology used to construct the comparison groups (see Tables IV.3 and IV.4). These results are especially striking for youth: not only do the magnitudes of the net-impact estimates for the control and the comparison group differ substantially, but also, for many of the estimates, the qualitative judgments about the impacts of the program differ between the control-group estimates and those based on the comparison groups.

Had the Supported Work evaluation for youth been based on any of the comparison-group methods experimented with in this study, net-impact estimates would have been obtained that differ qualitatively from the "true estimates" based on the control-group design actually used. Instead of concluding that the program led to short-run earnings gains and had negligible post-program effects, we probably would have concluded, based on the statistical match comparison group, that it had no effects for any time period, or, with any of the other comparison-group methods, that it had negative impacts in most time periods and no effects in others. Furthermore, in most cases, the apparent biases in the results are even stronger if either the difference in earnings model or a fixed-effects model is used rather than the basic earnings model.

For the AFDC sample, there is evidence of some upward bias in the estimates of net program impacts based on the comparison samples,

TABLE IV.3

ESTIMATED NET IMPACTS OF SUPPORTED WORK ON THE ANNUAL EARNINGS OF YOUTH: ALTERNATIVE MODELS AND COMPARISON GROUPS
(dollars per year)

Analytic Model	Supported Work Control Group	Random Samples		Cell Match Samples		Statistical Match Sample
		Stage 1 Screens	Stage 2 Screens	Short Match List	Long Match List	
1. Simple Difference in Earnings Gains						
1977 Earnings	326*	-427**	-382**	-319*	-794**	-176
1978 Earnings	-66	-1274**	-953**	-411*	-1018**	-607**
1979 Earnings	-7	-1982**	-1177	-617**	-1154**	-690**
2. Basic Earnings Model[a]						
1977 Earnings	313*	166	-624*	-388	-668*	-511
1978 Earnings	-28	-789	-939**	-699*	-1191**	-408
1979 Earnings	-18	-1937**	-890	-687*	-1179**	-339
3. Fixed-Effects Model						
Earliest Entry Year	-13	-346*	-405*	-286	-432*	-289*
Primary Entry Year	557**	-218	-271	-0	-259	36
Primary Entry Year + 1	110	-1151**	-1019**	-586**	-902**	-550**
Primary Entry Year + 2	-35	-1867**	-1340**	-715**	-1088**	-707**
Primary Entry Year + 3	-82	-2531**	-1529**	-753**	-1547**	-574*
Number of Observations						
Experimentals	69	69	69	69	69	69
Controls	87	158	321	320	112	71

NOTE: Standard deviations of the program impact estimates are presented in Appendix Table D.1.

[a] The full sets of regression coefficients for the models underlying these results for 1977 and 1979 are presented in Appendix Tables C.3 and C.4, respectively.

*Significant at 5 percent level.
**Significant at 1 percent level.

TABLE IV.4

ESTIMATED NET IMPACTS OF SUPPORTED WORK ON THE ANNUAL EARNINGS OF AFDC RECIPIENTS: ALTERNATIVE MODELS AND COMPARISON GROUPS
(dollars per year)

Analytic Model	Supported Work Control Group	Random Samples: Stage 1 Screens	Random Samples: Stage 2 Screens	Cell Match Samples: Short Match List	Cell Match Samples: Long Match List	Statistical Match Sample
1. Simple Difference in Earnings Gains						
1977 Earnings	1421**	1493**	1487**	1575**	1525**	1505**
1978 Earnings	431**	613**	676**	909**	695**	944**
1979 Earnings	285	540**	643**	843**	184	911**
2. Basic Earnings Model[a]						
1977 Earnings	1423**	1568**	1543**	1696**	1560**	1266**
1978 Earnings	505**	872**	472	782**	537	515
1979 Earnings	351*	710*	584	806**	257	683
3. Fixed-Effects Model						
Earliest Entry Year	-211	-236*	-336**	-211	176	-287*
Primary Entry Year	1243**	1225**	1092**	1317**	1330**	1181**
Primary Entry Year + 1	746**	785**	652**	959**	851**	752**
Primary Entry Year + 2	349**	498**	405**	777**	438*	487**
Primary Entry Year + 3	485*	792**	640**	1021**	594	687**
Number of Observations						
Experimentals	110	110	110	110	110	110
Controls/Comparisons	107	266	117	117	73	74

NOTE: Standard errors of the estimated net impacts are presented in Appendix Table D.2.

[a] The full sets of regression coefficients for the models underlying these results for 1977 and 1979 are presented in Appendix Tables C.5 and C.6, respectively.

*Significant at 5 percent level.
**Significant at 1 percent level.

especially in the simple earnings-gains model. However, the qualitative conclusions one would draw are roughly comparable across the analytic models and the comparison samples. The weight of the evidence in each case is that the program had large and significant impacts on earnings during the in-program period, and that, although the size of these impacts diminished over time, the point estimates based on comparison groups tended to be relatively large ($257 to $806 per year) as late as 1979.

For both the youth and the AFDC samples, the net-impact results estimated with the comparison groups vary considerably among the analytic models used. However, the nature of the effects of changing the model differ between the two samples. For the AFDC sample, the degree of apparent bias in the net-impact estimates is generally reduced by controlling for characteristics of the individual. However, for the youth sample, no clear pattern emerges as to which model yields the least biased estimates. One possible explanation is the fact that the basic earnings model is simply misspecified, and/or that the underlying behavioral models differ between the participant and comparison groups in ways that cannot be controlled for statistically. Although we could reject the similarity of the underlying models for the youth comparison and control samples in only a few instances, it is notable that the coefficients differ considerably both across samples and over time (see Appendix Tables A.7 and A.8), suggesting that the models may still differ.

The Use of Grouped Data. An important issue in this study is the extent to which the use of grouped data affects the study findings. Furthermore, since some types of sensitivity tests can most easily be conducted with ungrouped data, it is useful to examine whether the results

of analyses conducted with ungrouped data[1] provide useful approximations to the results based on fully grouped data. As indicated in Table IV.5, the impact estimates are reasonably similar whether they are based on grouped or ungrouped data.

Weighting. Two types of weights have been employed in this study. The first type is a correction for the heteroscedasticity introduced in a regression model by the use of grouped data. The second type is a weight on individual observations generated by the cell match procedure. For comparison groups in which individual observations are not weighted (random samples and statistical match samples), only the heteroscedasticity correction weight is relevant. In such cases, our regression results were obtained by applying weighted least squares to a model specified in terms of group means, where the weight for each group was the square root of the number of observations in that group (n_i). When one of the cell match comparison groups is used, our estimation technique accounted for both a heteroscedasticity correction weight and the cell match weights. In such cases, our regression results were obtained by applying weighted LS to a model specified in terms of weighted group means. In Appendix F, we demonstrate that the appropriate weight to apply to group i is:

$$(4) \qquad WT_i = \frac{\sum_j c_{ij}}{\sqrt{n_i}}$$

where c_{ij} is the cell match weight on observation j in group i.

For the long list cell match comparison groups, we tested the sensitivity of our results to various types of weighting specifications.

---

[1] Obviously, the outcome measures cannot be ungrouped for the Supported Work cases.

TABLE IV.5

NET IMPACT ESTIMATES USING GROUPED VERSUS UNGROUPED DATA WITH THE BASIC EARNINGS MODEL
(dollars per year)

Target Group	Comparison Sample and Type of Data			
	Supported Work Controls		Long List Cell Match Sample	
	Grouped Data	Ungrouped Data[a]	Grouped Data	Ungrouped Data
Youth				
1977 Earnings	313*	335**	-667*	-590**
1978 Earnings	-28	-47	-1191**	-1067**
1979 Earnings	-18	22	-1179	-1043**
AFDC Recipients				
1977 Earnings	1423**	1427**	1560**	1509**
1978 Earnings	505**	471**	537	848**
1979 Earnings	351*	321**	257	342*
Number of Observations				
Youth	156	1185	181	1590
AFDC Recipients	207	1404	183	1347

[a]All variables except the earnings measures are individual-specific rather than group mean values.

*Significant at 5 percent level.
**Significant at 1 percent level.

In Table IV.6 we present our results for four such specifications: (1) unweighted, (2) weighted for grouping only, (3) weighted for cell matching only, and (4) weighted for both cell matching and grouping. Results for the unweighted specification provide a baseline for comparison with the other three specifications. The unweighted results were obtained by applying (unweighted) LS to the model specified in terms of unweighted group means. Thus, each group was treated as if it were a single observation with a cell match weight of unity.

To weight for grouping only, we specified the model in terms of unweighted group means and applied weighted LS, where the weight for each group was $\sqrt{n_i}$. This is equivalent to assuming a cell-match weight of unity for each observation. In Table IV.6, we see that the results obtained for this specification are very similar to those obtained for the unweighted specification. This is to be expected since we have very little variation in group size (groups range in size from 7 to 10 observations). In general, the magnitude of the heteroscedasticity correction depends on the variation in group size, with such a correction being unnecessary when the the group size is constant.

To weight for cell matching only, we specified the model in terms of weighted group means and applied weighted LS, where the weight for each group was the sum of the cell match weights for all observations within that group. Thus, each group was treated as a single observation with a weight of $\sum_j c_{ij}$. This is consistent with setting n_i equal to unity for each group. In Table IV.6, we see that the results obtained for this specification are also very similar to those obtained for the unweighted

TABLE IV.6

NET IMPACT ESTIMATES GENERATED USING WEIGHTED VERSUS UNWEIGHTED REGRESSIONS FOR THE LONG LIST CELL MATCH COMPARISON GROUPS
(dollars per year)

	Target Group					
	Youth			AFDC Recipients		
	1977	1978	1979	1977	1978	1979
1. Unweighted	-676*	-1202**	-1513**	1441**	438	364
2. Weighted for Grouping Only	-669*	-1208**	-1504**	1427**	451	390
3. Weighted for Cell Matching Only	-656*	-1192**	-1151**	1567**	557	279
4. Weighted for Cell Matching and Grouping	-668*	-1191**	-1179**	1560**	537	257
Number of Observations						
Experimentals		69			110	
Comparisons		112			73	

NOTE: Appendix F discusses the weighting procedures used and the assumptions underlying them.

*Statistically significant at the 5 percent level.
**Statistically significant at the 1 percent level.

specification. This suggests that our model controls quite well for those factors that are related both to the matching criteria and to program status.

Finally, the fourth row of Table IV.6 contains our results when both heteroscedasticity correction weights and cell match weights are taken into account. For both the youth and AFDC samples, our results for this specification are also very similar to those obtained for the other three weighting specifications. We conclude that both types of weighting considered have minor impacts on our results.

Other Sensitivity Tests. For both the AFDC and the youth samples, but especially for the youth sample, we also conducted numerous other sensitivity tests and descriptive analyses in an effort to identify ways to improve the quality of the net-impact results obtained with the comparison groups, and to help us assess the generalizability of these findings to the CETA and JTPA populations. For the youth, where the comparison samples yield qualitatively and quantitatively different net-impact estimates than are obtained with the control group, we conducted net-impact analyses on sample subgroups in an effort to determine whether the greater flexibility introduced into the analytic model (permitting the equivalent of complete interaction between the control variables and the sample stratification characteristics) would lead to improved estimates, at least for some subgroups of youth. Among the sample stratifications experimented with were those based on interview earnings in the base year (none or some); Social Security earnings in the base year (low, medium, or high); years of school completed (fewer than 10 versus 10 or more); percent of time employed in the base year; central-city or non-central-city residence; and

gender. In no case were the net-impact estimates noticeably more reliable for any of these subgroups of youth than were the results for the overall youth population.

We also experimented with including a few additional control variables and variables that represent interactions of selected control variables into the youth equations. Again, however, the net-impact estimates tended to be insensitive to such changes in the model specification.

For the AFDC sample, we conducted only a few such sensitivity tests. The two primary tests entailed (1) excluding non-central-city residents from the comparison sample and (2) stratifying the samples by earnings in the base year. However, we also experimented with using a slightly different set of control variables. As was the case with the youth sample, we found that these changes in the samples and models had no effect on the overall conclusions about the quality of the comparison samples.

C. THE IMPLICATIONS OF THESE RESULTS FOR OTHER STUDIES THAT MUST RELY ON COMPARISON GROUPS

As was discussed in Chapter II, individuals in the Supported Work youth and AFDC samples are similar to CETA and JTPA participants in that they tend to have experienced severe employment problems. They differ in that both samples have less attachment to the work force than the typical CETA participant, and in that a higher proportion of the Supported Work youth sample exhibit characteristics that are associated with exceptionally low levels of employment and earnings--minority ethnic/racial composition, low educational levels, and limited employment experience. The key element

APPENDIX D

STANDARD ERRORS OF NET IMPACT ESTIMATES REPORTED IN CHAPTER IV